Capítulo

2

Interconexão de Redes na Internet do Futuro: Desafios e Soluções

Miguel Elias M. Campista, Lyno Henrique G. Ferraz, Igor M. Moraes,
Marcelo Luiz D. Lanza, Luís Henrique M. K. Costa e
Otto Carlos M. B. Duarte

GTA/PEE/COPPE/DEL-Poli - UFRJ - Rio de Janeiro, Brasil

Abstract

This short-course presents the main challenges for network interconnection in the Future Internet. We introduce the shortcomings of the Internet current model and the main proposals to solve them. Among all shortcomings, many are consequence of unforeseen demands by the Internet original design such as: mobility, multi-homing, multi-path, and network scalability. These challenges have been attracting many research efforts in the last few years because of their relevance and complexity. In this short-course, new protocols for network interconnection are presented, from IP extensions to radical proposals to replace IP in the Future Internet. In addition, we introduce new intra- and interdomain routing protocols as well as experimental results obtained with one of these proposals prototypes.

Resumo

Este minicurso apresenta os principais desafios para a camada de interconexão de redes da Internet do Futuro. As limitações do modelo atual da Internet e as principais propostas para solucioná-las são apresentadas. Dentre os desafios, muitos são consequências de demandas não previstas pelo projeto inicial da Internet, como: mobilidade, redes multidomiciliadas, múltiplos caminhos e escalabilidade da rede. Esses desafios têm atraído muito investimento em pesquisa nos últimos anos devido à relevância e à complexidade do tema. Neste minicurso, são apresentados novos protocolos para interconexão de redes, desde extensões do IP até propostas radicais para substituir o IP na Internet do Futuro. Além disso, propostas para novos protocolos de roteamento intra e interdomínio são apresentadas, bem como resultados práticos obtidos com um dos protótipos dessas propostas.

2.1. Introdução

No início dos anos 80, a Internet surgia com a adoção da pilha de protocolos TCP/IP. Naquela época, a Internet era composta por estações fixas de grande porte cujas principais aplicações eram o acesso remoto e a troca de arquivos. Nesse cenário, o IP (*Internet Protocol*) tinha como objetivo atender requisitos fundamentais como interconectar redes, prover conectividade fim-a-fim e garantir acessibilidade global [Clark et al. 2004]. Esses requisitos deviam ser atendidos seja qual fosse a aplicação e a tecnologia de controle de acesso ao meio utilizada. Além do IP, os protocolos de roteamento também eram utilizados na interconexão de redes. Esses protocolos cumpriam o mesmo papel dos protocolos de roteamento atuais. Entretanto, eles podiam calcular os menores caminhos entre qualquer par origem-destino, já que a Internet era composta por poucos nós e era apenas subdividida em poucas redes de instituições sem fins lucrativos. Cada roteador, por conseguinte, conhecia o melhor caminho disponível para qualquer uma das redes da Internet. Um exemplo disso, foi o projeto do GGP (*Gateway-to-Gateway Protocol*) [Hinden e Sheltzer 1982] que foi um dos primeiros protocolos de interconexão de redes. O GGP listava em suas mensagens de atualização as distâncias em número de saltos para todas as redes da Internet e suportava até no máximo 256 redes.

Nos anos seguintes, o número de usuários e de redes na Internet aumentou aceleradamente. Esse aumento tornou as atualizações de topologia mais frequentes e as atualizações do GGP mais complexas, visto que cada uma das redes era administrada por grupos ou instituições diferentes. Além disso, havia mais de uma versão do GGP em uso, o que dificultava a interoperação das redes. Essas dificuldades levaram a Internet a ser dividida em Sistemas Autônomos (*Autonomous Systems* - ASes), definidos como um conjunto de redes e roteadores administrados por um grupo ou uma instituição comum. Cada Sistema Autônomo escolhe o seu próprio protocolo de roteamento interno, chamado protocolo intradomínio, e para se comunicar com redes em outros Sistemas Autônomos, utiliza um protocolo interdomínio. O protocolo interdomínio era comum a todos os ASes já que as informações de roteamento precisam ser trocadas para manutenção global da conectividade. O primeiro protocolo interdomínio desenvolvido foi o EGP (*Exterior Gateway Protocol*) [Mills 1984] que também sucumbiu devido a problemas de escalabilidade. O EGP foi substituído pelo BGP (*Border Gateway Protocol*) [Lougheed e Rekhter 1989] que se tornou o protocolo interdomínio da Internet.

A organização da Internet de hoje é a mesma que surgiu após a sua divisão em Sistemas Autônomos. Além disso, a Internet ainda deve atender todos os seus requisitos fundamentais, pois foi por isso mesmo que se tornou o sucesso de hoje. Outra característica mantida foi o crescimento do número de ASes e, consequentemente, de redes. Dados mostram que a Internet atual consiste de 7,4 milhões de redes compostas por no máximo 254 estações (redes /24). Essas redes representam 95% de todas as redes alcançáveis da Internet [Caida 2010]. Outros dados mostram que o crescimento da Internet foi de 380,3% de 2000 até 2009 [Internet World Stats 2010]. Tal crescimento é impulsionado pela oportunidade de negócios e pela popularização da Internet. A oportunidade de negócios surgiu da necessidade de trânsito de dados entre ASes não conectados diretamente. Essa necessidade levou a acordos entre ASes vizinhos que estão no caminho entre pares origem-destino para transporte de dados. Já a popularização da Internet é consequência da facilidade e rapidez na obtenção de informações proporcionadas para

usuários corporativos e residenciais que pagam os provedores de acesso pelo serviço.

Assim como ocorreu no passado, mais uma vez a Internet está chegando a um ponto onde mudanças se tornam essenciais. O sucesso da Internet não vem sem consequências. Mesmo os protocolos atuais como o BGP e o próprio IP vêm apresentando sinais de fadiga. Um indicativo disso é o BGP ter sido proposto em 1989 e ter sido atualizado até a sua versão quatro em apenas seis anos [Lougheed e Rekhter 1989, Rekhter e Li 1995]. Em 2006, o BGP-4 foi atualizado novamente para corrigir erros de edição de RFCs (*Request For Comments*) anteriores [Rekhter et al. 2006]. Embora o BGP seja fundamental, os problemas dele também são oriundos das próprias limitações do IP, que não acompanhou no mesmo passo a evolução da Internet. Portanto, além dos problemas relacionados ao crescimento da Internet, nem o projeto inicial do IP nem tampouco dos protocolos de roteamento existentes foram desenvolvidos tendo em vista a possibilidade de evoluções para atender demandas emergentes. Dentre essas demandas, as mais relevantes são o suporte à mobilidade, o suporte aos múltiplos domicílios (*multi-homing*), o suporte aos múltiplos caminhos e a escalabilidade dos roteadores. Embora novas propostas ou extensões tenham surgido como resposta às novas demandas, p. ex. o IPv6 [Deering e Hinden 1998], o IP móvel [Perkins 2002] e o multicast [Deering 1988, Costa et al. 2006], nenhuma delas solucionou o problema por completo, e em alguns casos, nem foram aplicadas na prática.

Atualmente, investiga-se qual a melhor abordagem a ser tomada naquela que será a Internet do Futuro. Dentre as possíveis, estão a ruptura total com o modelo atual da Internet e o consequente surgimento de novos protocolos para cumprir um papel semelhante ao do IP [Crowcroft 2008] ou a continuidade da elaboração de novas extensões aos protocolos [Ratnasamy et al. 2005]. A primeira abordagem pode ir de encontro aos interesses dos administradores dos ASes que já possuem base instalada operacional. Já a segunda abordagem pode apenas adiar o problema enfrentado atualmente. Um ponto importante é a possibilidade de novos protocolos incluírem todas as novas demandas emergentes ou deixarem ainda algumas como casos a parte. Em comum, pode-se destacar que a nova proposta deverá ser escalável e capaz de evoluir para atender os futuros avanços da Internet. Este capítulo apresenta os principais desafios em interconexão de redes na Internet do Futuro, bem como as limitações do modelo atual e as propostas para solucioná-las. Adicionalmente, alguns dos novos protocolos de roteamento assim como resultados experimentais são apresentados. Este capítulo está organizado da seguinte maneira. A Seção 2.2 apresenta a arquitetura atual da Internet e as definições empregadas neste capítulo. A Seção 2.3 discute os principais desafios enfrentados na Internet atual para prover serviços como: mobilidade, múltiplos domicílios, múltiplos caminhos, caminhos programáveis e escalabilidade dos roteadores. Essa seção ainda apresenta as limitações da arquitetura atual que impedem que tais serviços sejam oferecidos plenamente. A Seção 2.4 descreve os novos conceitos da área para aprimorar a interconexão de redes na Internet do Futuro, bem como suas principais propostas. A Seção 2.5 apresenta resultados experimentais obtidos com protótipos de uma das novas propostas para a Internet do Futuro. Por fim, a Seção 2.6 conclui este capítulo.

2.2. Arquitetura Atual da Internet

A arquitetura atual da Internet é composta por diferentes redes interconectadas através do IP. As redes administradas pela mesma instituição são denominadas Sistemas Autônomos (ASes). Cada AS executa um protocolo de roteamento intradomínio para interconectar seus próprios roteadores enquanto os diferentes ASes executam um protocolo de roteamento interdomínio para se comunicarem. Essa arquitetura é uma evolução da arquitetura original da Internet que era apenas subdividida em redes de instituições diferentes. A Figura 2.1 mostra a topologia da Internet no início dos anos 80 [History Museum 2010].

A posição de um AS na topologia da Internet pode ser usada para classificação. Um AS diretamente conectado aos usuários é denominado de AS de borda ou provedor de acesso. Tais ASes tarifam os usuários pelo acesso à Internet. Os usuários normalmente estão localizados em redes de acesso, que também podem ser denominadas de redes *stub*. As redes *stub* conhecem apenas um caminho para outras redes. Esses caminhos são chamados de rotas *default* já que são sempre utilizadas para alcançar qualquer outra rede. As rotas *default* são anunciadas para as estações da rede de acesso pelo ponto de interconexão (*gateway*) da rede. Uma rede de acesso recebe uma faixa de endereços IP pertencente à faixa do AS de borda. Essa faixa é utilizada para atribuição de endereços às estações da rede. Logo, o prefixo comum dos endereços IP das estações pode ser usado para identificação da rede. Da mesma maneira, um endereço IP identifica uma estação.

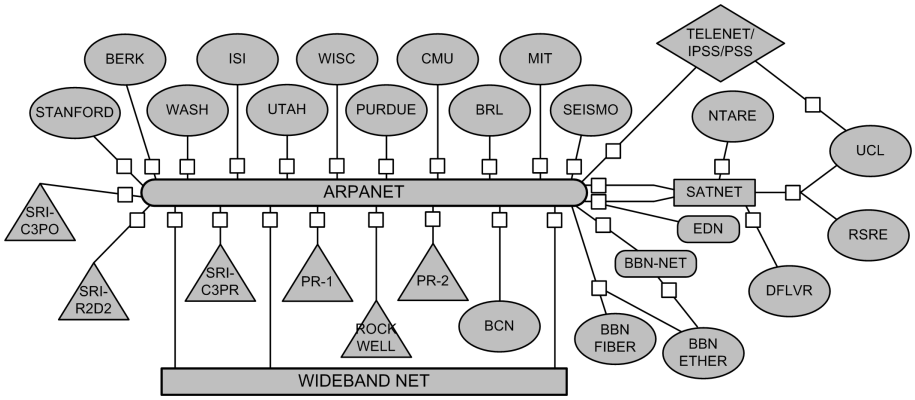


Figura 2.1. A Internet no início do anos 80.

Os ASes que não provêem acesso diretamente a usuários formam o núcleo da Internet. Esses ASes são conhecidos como ASes de trânsito por oferecerem serviços de transferência de dados entre ASes. Os ASes de trânsito são responsáveis por encaminhar todo o tráfego entre as bordas da Internet. Para isso, assume-se que toda origem e destino de tráfego estão localizados na borda da rede. Essa premissa está de acordo com um dos requisitos da Internet que é a manutenção da inteligência nas bordas para que o núcleo seja simples. Os ASes de trânsito tarifam os seus ASes vizinhos, sejam eles de borda ou outros de trânsito, dependendo do tipo de acordo estabelecido e da quantidade de tráfego injetado pelo vizinho no AS de trânsito.

As instituições responsáveis pelos ASes de trânsito são denominadas de provedores de serviço da Internet (*Internet Service Providers* - ISPs). Os ASes de trânsito normal-

mente não possuem rotas *default*. Para isso, assume-se que os ISPs possuem visão global de todos os ASes da Internet. A ausência de rotas *default* rendeu ao conjunto de ASes de trânsito o nome de zona livre de rota *default* (*Default Free Zone* - DFZ). A Figura 2.2 mostra a arquitetura atual da Internet e as suas subdivisões em Sistemas Autônomos (ASes). Os ASes de borda possuem ligadas a eles as redes de acesso dos usuários.

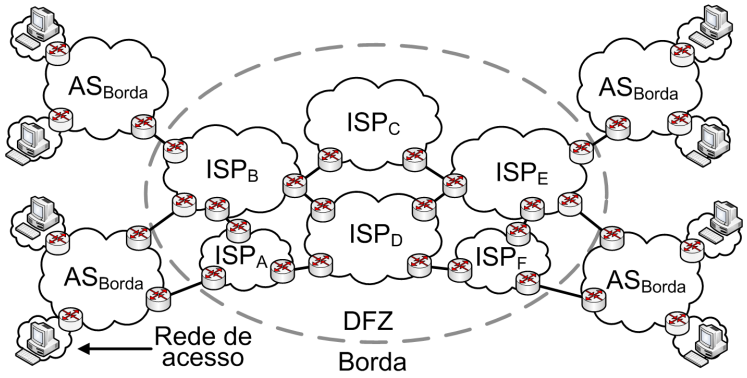


Figura 2.2. A arquitetura da Internet atual.

Duas definições importantes que serão usadas posteriormente são a FIB (*Forwarding Information Base*) e a RIB (*Routing Information Base*). A FIB e a RIB são tabelas com funções distintas utilizadas pelos roteadores da Internet. A FIB é a tabela de roteamento que é consultada sempre que o roteador recebe um pacote. Após uma consulta à tabela de roteamento, o roteador encaminha o pacote recebido pela interface de saída correspondente ao caminho escolhido para alcançar a rede de destino. A RIB, em contrapartida, é a base de dados que é utilizada para construir a FIB. Por exemplo, no protocolo OSPF (*Open Shortest Path First*), a FIB contém o mapa conhecido da topologia da rede. As FIBs são construídas para acelerar a consulta a interfaces de saída enquanto as RIBs são construídas para tornar a atualização da base de dados mais rápida. A Figura 2.3 ilustra um esquema básico de um roteador. A figura mostra o caminho percorrido pelos pacotes de controle (caminho de controle) e pelos pacotes de dados (caminho de dados).

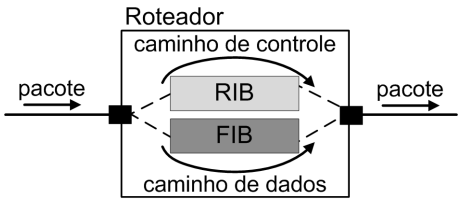


Figura 2.3. Esquema básico de um roteador atual.

2.3. Desafios em Interconexão de Redes

Muitos dos desafios enfrentados pela Internet atual estão diretamente ligados à sua arquitetura e ao crescimento acelerado do número de estações e redes. A arquitetura original da Internet não previu o suporte a serviços como mobilidade, múltiplos domicílios, múltiplos caminhos e programabilidade de caminhos. Esta seção discute os principais desafios da interconexão de redes e do roteamento na Internet do futuro.

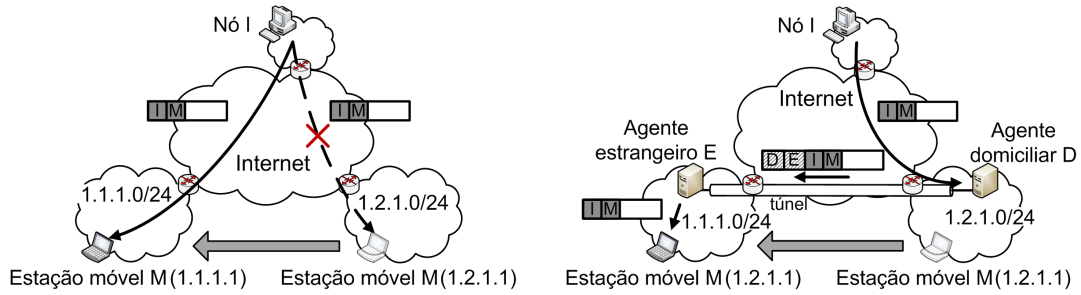
Mobilidade

A disseminação das redes sem-fio é um dos principais desafios à arquitetura original da Internet. As redes sem-fio possibilitam que as estações se movimentem. Logo, uma vez que a arquitetura original da Internet não previa a mobilidade, muitas das soluções utilizadas não se adequam a essa característica. Um exemplo típico é o funcionamento do TCP (*Transmission Control Protocol*) em redes sem-fio [Hanbali et al. 2005]. O TCP foi desenvolvido para uso na Internet cabeada. Por isso, esse protocolo considera que atrasos na recepção de reconhecimentos positivos ou perda de pacotes são devido a congestionamentos na rede já que as perdas por erro de transmissão são negligenciáveis. Assim, para evitar o colapso da rede, as fontes TCP reduzem as suas taxas de transmissão quando detectam perdas. Nas redes sem-fio essa mesma premissa pode não ser mais verdadeira visto que as perdas por erro de transmissão são consideráveis. Portanto, reduzir a taxa de transmissão representa um desperdício de banda passante uma vez que a rede não enfrenta problemas de congestionamento.

A mobilidade das estações introduz desafios para camada de interconexão de redes, em especial no que concerne o endereçamento e o roteamento. Na Internet original, o endereçamento foi organizado de maneira hierárquica para possibilitar a agregação de rotas e assim aumentar a escalabilidade do roteamento. Um exemplo dessa organização é visto nos prefixos usados pelas redes de acesso ligadas a um AS. Os prefixos das redes são faixas pertencentes ao AS de borda. Assim, o AS de borda pode anunciar aos outros ASes vizinhos os prefixos agregados das redes ligadas a ele. A consequência da agregação é a redução do número de entradas nas tabelas de roteamento e a simplificação da busca nas FIBs para encaminhar pacotes às estações.

A estrutura hierárquica da Internet levou à organização geográfica de endereços IP. Tal organização associou ao endereço IP a localização da estação na Internet. Por outro lado, a identificação de estações da Internet é feita através de nomes que são mapeados pelo DNS (*Domain Name System*) em endereços IP. Assim, o endereço IP é utilizado para localização e identificação de uma estação, o que caracteriza a "semântica sobrecarregada" do endereço IP. Esse problema representa um entrave à mobilidade das estações já que ao se moverem, as estações mudam a sua localização e, consequentemente, devem mudar o seu endereço IP. Porém, a arquitetura da Internet atual assume que os endereços IP são invariantes durante toda a comunicação entre estações. Como padrão, se uma estação muda de rede de acesso à Internet, essa estação deve reconfigurar o seu endereço IP para um que seja topologicamente correto de acordo com a sua nova rede. Ao receber um novo endereço IP, as comunicações estabelecidas com o endereço IP anterior são perdidas. Além disso, o estabelecimento de conexão com uma estação que muda constantemente de endereço dificulta a resolução de nomes já que a estação não possui um endereço permanente. A Figura 2.4(a) ilustra a reconfiguração do endereço IP da estação móvel M ao mudar de rede de acesso. A mudança do endereço provoca o restabelecimento da conexão com o nó I da Internet.

Soluções adicionais ao IP devem ser propostas para que as comunicações, e possivelmente as conexões pré-estabelecidas, sejam mantidas mesmo após o deslocamento das estações. Uma proposta inicial para resolver o problema foi o IP móvel [Perkins 2002]. O IP móvel mantém as conexões de uma estação móvel através da manutenção da associação



(a) Mudança de endereço em decorrência da mudança de rede de acesso.

(b) Funcionamento do IP móvel.

Figura 2.4. Mobilidade de estação.

dessa estação com o seu endereço IP original, mesmo após o deslocamento. Para isso, é necessário que a estação móvel informe à rede original qual a sua nova localização para que todo pacote recebido possa ser encaminhado para a estação em seu novo endereço na rede estrangeira. A nova localização da estação móvel é recebida pelo agente domiciliar (*home agent*), que é uma estação na rede de acesso original responsável por receber e encaminhar todo o tráfego destinado à estação móvel. O agente domiciliar encapsula os pacotes destinados à estação móvel e os encaminha através de um túnel estabelecido entre o agente domiciliar e uma estação da rede estrangeira. Essa estação, denominada agente estrangeiro (*foreign agent*), é responsável por atribuir um endereço IP válido da rede estrangeira à estação móvel e também é responsável por encaminhar os pacotes recebidos através do túnel até essa mesma estação. A estação móvel recebe um endereço IP válido do agente estrangeiro após um procedimento de registro. O procedimento de encaminhamento, ilustrado pela Figura 2.4(b), pode ser resumido da seguinte maneira. Todos os pacotes enviados à estação móvel M são recebidos pelo agente domiciliar D. Esse agente encapsula os pacotes e os encaminha pelo túnel até o agente estrangeiro E. O agente estrangeiro E desencapsula os pacotes e os encaminha até a estação móvel M. O tráfego no sentido reverso é originado pela estação móvel M e enviado até o agente estrangeiro E. Esse encapsula os pacotes e os envia até o agente domiciliar D através do túnel. O agente domiciliar D recebe os pacotes, os desencapsula e os envia para a Internet. Apesar de o procedimento permitir o deslocamento das estações móveis, o IP móvel encaminha os pacotes de maneira indireta entre as estações de origem e destino. Esse encaminhamento indireto é refletido em um caminho mais longo percorrido pelos pacotes, o que reduz a eficiência do roteamento na Internet.

O IPv6 [Johnson et al. 2004] móvel dispensa o uso do agente estrangeiro. Portanto, o túnel é estabelecido entre o agente domiciliar e a própria estação móvel. Além disso, o IPv6 móvel permite que os pacotes sejam encaminhados da Internet até a estação móvel sem precisar passar pela rede domiciliar. O encaminhamento direto e possível com o uso de uma extensão de cabeçalho para roteamento (*routing header extension*) do IPv6. A extensão permite que a estação móvel utilize o endereço IP da rede estrangeira como endereço de origem e liste na extensão do cabeçalho o endereço da rede domiciliar. O nó da Internet ao receber o pacote reconhece o novo endereço da estação e, após um procedimento de associação, utiliza o endereço da estação móvel como endereço de destino dos próximos pacotes. Mesmo com essas diferenças para o IP móvel, o IPv6 móvel ainda

exibe problemas como latência no restabelecimento do acesso à Internet, principalmente se a estação móvel se deslocar a altas velocidades.

Múltiplos domicílios

Os múltiplos domicílios (*multi-homing*) são redes de acesso ou estações que se conectam à Internet através de múltiplas conexões. Essas conexões são obtidas de diferentes ISPs ou do mesmo ISP, mas usando faixas de endereços diferentes. A Figura 2.5 ilustra uma estação multidomiciliada (*host multihoming*) e uma rede multidomiciliada (*site multihoming*). No primeiro caso, a estação possui interfaces configuradas com endereços IP diferentes. Já no segundo caso, o ponto de interconexão possui saídas para dois ISPs diferentes, porém as estações internas podem possuir endereços IP privados ou podem ter a sua única interface configurada com endereços IP de cada um dos ISPs. Assim, as redes multidomiciliadas podem ser identificadas por mais de uma faixa de endereços, que devem ser anunciadas para todos os ASes da Internet através de rotas do BGP. A motivação para o emprego dos múltiplos domicílios é o aumento da confiabilidade na comunicação com a Internet [Farinacci et al. 2009b, Nordmark e Bagnulo 2009]. Em caso de falha de um dos provedores, ainda há alternativas para a manutenção da comunicação. Para maior confiabilidade, as redes multidomiciliadas devem estar configuradas de maneira a isolar eventuais falhas individuais. Na ausência de falhas, as múltiplas faixas de endereços podem ser também exploradas para possibilitar engenharia de tráfego, maximizar vazão agregada e até mesmo reduzir custos. Um exemplo de redução de custos é o uso de um ISP mais caro para tráfego que exige qualidade de serviço e o uso de um mais barato, no caso contrário.

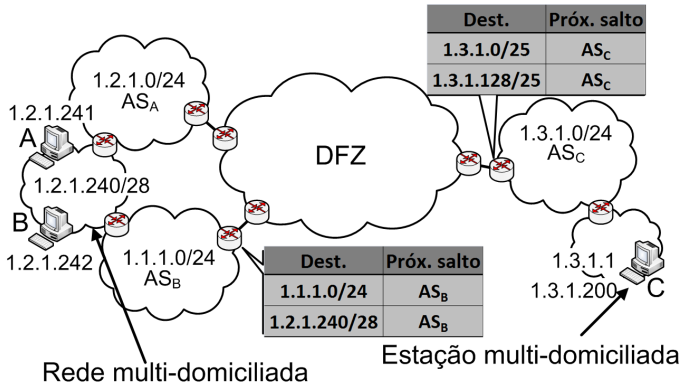


Figura 2.5. Exemplos de estação e rede multidomiciliada.

A utilização dos múltiplos domicílios traz dificuldades do ponto de vista do roteamento. As múltiplas faixas podem ser alocadas por ISPs (*provider assigned*) ou podem ser independentes do ISP (*provider independent*). Nessa última opção, a faixa de endereços é alocada diretamente pelo órgão regional de distribuição de endereços da Internet. A principal vantagem em se utilizar faixas independentes é evitar reconfiguração caso haja a troca de provedor de acesso. O funcionamento dos múltiplos domicílios viola a organização hierárquica de endereços da Internet, o que impede a agregação de endereços. Essa técnica requer o anúncio das múltiplas faixas desagregadas para que todas as outras redes na Internet tomem conhecimento da localização daquela faixa. Uma vez que cada

provedor de serviço possui suas faixas de endereços pré-estabelecidas, anunciar faixas de endereços desagregadas ou de outros provedores causa problemas de escalabilidade, como o aumento das tabelas BGP.

Múltiplos caminhos

Muitos protocolos de roteamento atuais escolhem apenas um único caminho para interconectar um par origem-destino na rede. Protocolos intradomínio como o OSPF e o RIP (*Routing Information Protocol*) tipicamente mantêm em suas tabelas de roteamento apenas o vizinho que pertence ao melhor caminho até o destino. O OSPF em especial oferece variantes que calculam múltiplos caminhos, mas que são pouco exploradas na prática. Essa limitação dificulta a introdução de confiabilidade e flexibilidade no roteamento, uma vez que seria possível dividir o tráfego entre caminhos diferentes. Outras vantagens são a possibilidade de escolher caminhos de acordo com os requisitos das aplicações, por exemplo, caminhos com maior banda ou menor atraso, e ainda evitar caminhos congestionados [He e Rexford 2008].

Os múltiplos caminhos representam uma opção pouco explorada na Internet apesar do seu potencial. Medidas demonstram que existe um caminho alternativo com taxa de perda e atraso menores que o caminho usado em 30 a 80% do tempo [He e Rexford 2008]. O emprego desses caminhos pode aumentar significativamente o desempenho do encaminhamento de pacotes na Internet. Entretanto, os múltiplos caminhos deixam de ser explorados porque muitas vezes esbarram em limitações de camadas superiores. Um exemplo disso é a possibilidade do desordenamento de pacotes de uma mesma conexão TCP se a diversidade de caminhos for explorada no nível de pacotes e não no nível de fluxos. A Figura 2.6 ilustra o exemplo do uso dos múltiplos caminhos entre as estações A e C. Os múltiplos caminhos são usados para envio de tráfego ao mesmo tempo, o que difere essa estratégia do uso dos múltiplos domicílios. A Figura 2.6 ilustra um exemplo de múltiplos caminhos totalmente disjuntos. Entretanto, os múltiplos caminhos podem ser parcialmente disjuntos, no caso de alguma rede de acesso ou AS que não ofereça múltiplas saídas para a Internet, por exemplo.

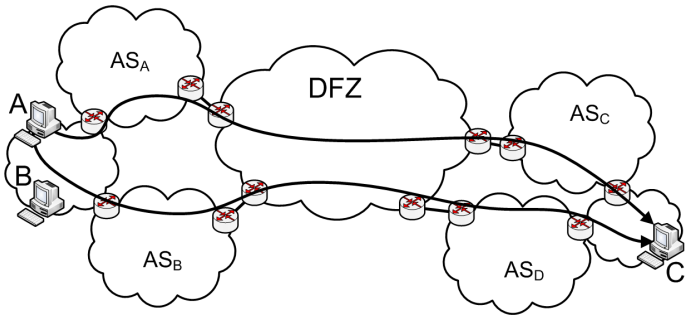


Figura 2.6. Exemplos do uso dos múltiplos caminhos na Internet.

Embora a topologia da Internet ofereça múltiplos caminhos devido às múltiplas conexões entre ASes e aos múltiplos caminhos intradomínio, problemas relacionados com escalabilidade, acordos comerciais entre os diferentes ISPs e o próprio projeto do BGP dificultam o emprego dessa possibilidade [He e Rexford 2008]. A primeira questão é uma

consequência da obrigação dos roteadores em armazenarem informações de múltiplos caminhos para cada destino. Essa característica exige a execução de algoritmos ou variações de algoritmos com complexidade superior ao Dijkstra ou ao Bellman-Ford, usados em muitos dos protocolos intradomínio, que aumenta a troca de informações de controle e ainda aumenta o número de entradas nas tabelas de roteamento [He e Rexford 2008]. A segunda questão é consequência da divisão da Internet em múltiplos ASes. Para serem plenamente explorados, os múltiplos caminhos devem ter condições de passar por redes que pertençam a diferentes organizações. Tais organizações não necessariamente possuem acordos entre si e, portanto, não encaminham tráfego vindo desses ASes. Já a última questão está relacionada com a maneira como o BGP foi desenvolvido.

O BGP, assim como os protocolos intradomínio, foi desenvolvido para ser um protocolo que ofereça apenas uma opção de caminho para cada prefixo de rede da Internet. Cada ISP anuncia aos seus vizinhos um único caminho ativo para alcançar outras redes através dele. O caminho disponibilizado depende das políticas utilizadas pelo BGP. Entretanto, se mais de um caminho fosse oferecido pelos ISPs aos seus vizinhos, parte do controle do tráfego encaminhado seria perdido. Isso ocorre pois não há como prever a distribuição de tráfego realizada pelos ISPs vizinhos que é encaminhado através das múltiplas opções oferecidas. Esse controle é importante para que os ISPs possam realizar engenharia de tráfego dentro da sua própria rede. Algumas opções para forçar os múltiplos caminhos sem a necessidade de cada ISP anunciar múltiplos caminhos é através do uso do roteamento pela fonte ou das redes sobrecamada (*overlay*). Em ambas as opções, o ISP também perde parte do controle de sua rede. Isso porque tanto a fonte que escolhe o caminho quanto as redes formadas em camadas superiores podem desconsiderar as políticas de cada ISP e as ligações na camada física, respectivamente.

Caminhos programáveis

Atualmente, os caminhos seguidos pelos pacotes na Internet são definidos por protocolos de roteamento que, em especial o BGP, envolvem muitas configurações manuais de administradores de rede. Os protocolos de roteamento intradomínio escolhem o melhor caminho entre a origem e o destino baseado em métricas estabelecidas, p. ex., o número de roteadores atravessados, e os protocolos de roteamento interdomínio consideram acordos comerciais para escolher qual dos ISPs vizinhos deve ser utilizado. Nesse último caso, as condições nas quais o tráfego é encaminhado dependem do tipo de acordo entre as partes que pode inclusive limitar requisitos de rede como a banda passante máxima.

Uma possibilidade para conceder maiores poderes aos usuários é criar critérios para que os caminhos escolhidos na Internet sejam programáveis ou configuráveis de acordo com requisitos desejados pelos próprios usuários ou por aplicações no nível de usuário. Tais requisitos podem ser demandas de uma aplicação, como por exemplo, garantias de qualidade de serviço, que poderiam ser escolhidas dinamicamente por agentes inteligentes ou pelos usuários [Clark et al. 2004]. Para isso, os caminhos da Internet não devem se basear somente em métricas de interesse dos provedores de serviço ou em acordos comerciais, mas também em métricas de interesse dos usuários. Essa possibilidade, porém, esbarra no serviço atualmente utilizado na Internet que é o serviço de "melhor esforço" oferecido pelo IP. Para possibilitar a programação de caminhos a Internet deve suportar qualidade de serviço e a possibilidade dos usuários ou agentes fazerem escolhas

que podem ir contra os acordos dos provedores. Além disso, uma das características fundamentais da Internet é a manutenção da inteligência nas bordas da rede. Essa premissa é baseada no fato que não há ninguém melhor do que os usuários para saber se a aplicação possui bom desempenho ou não [Clark et al. 2004]. Portanto, requisitos como confiabilidade devem ser verificados nas bordas da rede e qualquer outro tipo de verificação no núcleo é considerado redundante. Essa premissa é satisfeita se os caminhos forem escolhidos pelos usuários. Entretanto, a possibilidade do emprego de agentes pode ir em direção oposta a essa premissa já que os agentes podem ser inseridos no núcleo da rede.

A liberdade de escolha de caminhos pelos usuários pode incentivar a disputa e, assim, provocar a redução dos custos de acesso. Uma questão em aberto é a possibilidade da ausência da qualidade de serviço fim-a-fim ser apenas uma consequência da falta de competição entre os ISPs. Caso fosse possível escolher caminhos, usuários que possuíssem aplicações com requisitos especiais poderiam optar por caminhos mais apropriados. Nesse caso, o ISP que entrasse no mercado oferecendo essa possibilidade levaria vantagem sobre os outros que seriam obrigados a modificar os seus serviços para serem competitivos [Clark et al. 2004]. Caso nenhum ISP tomasse a iniciativa, incentivos poderiam ser oferecidos aos ISPs para que esses começassem a oferecer a possibilidade de escolha aos usuários. O problema de oferecer liberdade aos usuários ou agentes é o aumento da complexidade que cada um deve lidar. Por exemplo, os usuários ou agentes precisariam de conhecimento global da rede e ainda poderiam introduzir falhas causadas por más configurações ou más opções. Além dos problemas mencionados, a possibilidade de escolha de caminho pode exigir que cada nó mantenha todos os possíveis caminhos em memória. Esse requisito pode ser mais uma fonte de problemas de escalabilidade.

Escalabilidade dos roteadores

Um dos maiores desafios do aumento acelerado do número de estações e redes na Internet é a escalabilidade. De acordo com um recente relatório do IAB (*Internet Architecture Board*), a escalabilidade é um dos desafios mais críticos em curto prazo para o projeto da Internet do Futuro [Meyer et al. 2007]. Os roteadores são equipamentos que possuem recursos limitados de memória e processamento. Portanto, o aumento indefinido do tamanho das tabelas de roteamento (FIB) e das bases de dados (RIB) pode afetar o desempenho do encaminhamento de pacotes. Cada roteador dispõe de poucos instantes para armazenar um pacote, selecionar a interface de saída baseado no prefixo de rede do endereço de destino e encaminhar para interface correspondente. Além disso, a sobrecarga de controle gerada ao aumentar o número de estações pode se tornar relevante diante da banda passante disponível. Para exemplificar tal crescimento, dados do CIDR (*Classless Inter-Domain Routing*) mostram que desde o início da década de 90 até os dias atuais o número de entradas BGP ativas nas FIBs aumentou de algumas unidades para aproximadamente 310 mil [CIDR 2010].

Os desafios ligados à escalabilidade poderiam ter menor impacto caso o projeto original da Internet fosse seguido. A premissa de organização hierárquica da Internet e a consequente agregação de endereços têm como objetivo tornar o roteamento escalável [Massey et al. 2007, Jen et al. 2008]. Entretanto, as demandas emergentes discutidas nesta seção impedem a agregação de endereços nas FIBs e/ou aumentam a carga de controle da rede. O suporte às redes multidomiciliadas e aos múltiplos caminhos reque-

rem a associação de prefixos distintos a uma mesma rede e requerem a associação de mais de um caminho para um mesmo destino. Já o suporte à mobilidade aumenta a quantidade de informações de controle trocadas para manter uma estação móvel conectada e destrói a premissa de organização hierárquica da Internet. Por fim, a programação de caminhos resulta em armazenamento de informações do roteamento nos usuários. Um efeito indireto que também influencia na escalabilidade é o uso de espaços de endereçamento maiores que o IPv4, como é o caso do IPv6. O uso do IPv6 pode aumentar a oferta de endereços IP, o que resulta em mais estações e redes e mais carga de controle.

Os problemas de escalabilidade já têm sido observados na prática pelo tamanho das tabelas de roteamento BGP. Essas tabelas têm aumentado devido à agregação parcial das faixas de endereços anunciadas pelos ISPs. A Figura 2.5 mostra que o roteador de borda do AS_B anuncia um prefixo de rede que não é agregável ao seu. Já o roteador de borda do AS_C anuncia o seu prefixo desagregado para possibilitar a estação multidomiciliada. Essa agregação parcial é muitas vezes interessante para um provedor de serviço deixar de agregar determinadas faixas de endereços para inclusive balancear carga ou para oferecer serviços diferenciados para os clientes. Dados do CIDR mostram que a razão entre redes que sofreram agregação de prefixos e o total de redes na Internet é atualmente de 38,4% [CIDR 2009], o que demonstra a relevância do problema na Internet. A agregação parcial de endereços, assim como muitos dos desafios da Internet do Futuro, pode culminar no aumento das tabelas de roteamento e no aumento de sobrecarga de controle. Essas consequências representam um entrave ao aumento do número de estações já que a busca por endereços de destino pelos roteadores e o processamento de mensagens podem ser operações custosas computacionalmente.

2.4. Propostas para a Interconexão de Redes

Nesta seção são apresentados novos conceitos e novas propostas que têm por objetivo resolver os desafios destacados na seção anterior. Muitas dessas propostas rompem completamente com a arquitetura original da Internet e algumas delas levam em conta um período de transição para a sua implementação.

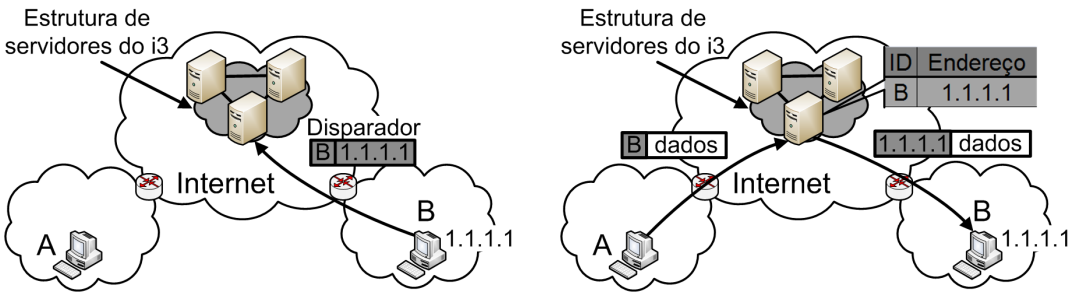
2.4.1. Separação de Localização e Identificação

Uma das principais propostas para minimizar os problemas de escalabilidade devido às demandas por mobilidade e múltiplos domicílios na Internet é a separação da localização física da identificação das estações. Essa separação tem o intuito de quebrar a semântica sobrecarregada do IP e, por isso mesmo, vem sendo considerada como fundamental para a Internet do Futuro [Caesar et al. 2006b]. Esse tipo de abordagem é denominada *Loc/ID split*. O *Loc/ID split* permite que as comunicações realizadas a partir de um identificador invariável se mantenham, mesmo que a localização da estação mude. Ainda que em outro nível, o procedimento é semelhante ao DNS que separa a identificação do nome de um sítio da Internet do seu endereço IP. Entretanto, assim como no caso do DNS, é também necessário algum sistema de mapeamento para associar identificador e localizador. O mapeamento dinâmico e escalável entre identificador e localizador é um problema de pesquisa em aberto [Iannone e Bonaventure 2007, Luo et al. 2009].

O i3 (*Internet Indirection Infrastructure*) [Stoica et al. 2004] baseou sua proposta

na constatação de que serviços como mobilidade e multicast endereçam estações de maneira indireta para funcionarem na Internet atual. Por exemplo, para prover suporte à mobilidade há possivelmente o emprego de agentes domiciliares assim como para prover suporte ao multicast há o emprego de endereço de grupo. Em ambos os casos o endereço de destino utilizado pela fonte não é o endereço da estação final. O i3 utiliza essa constatação para propor uma estrutura sobrecamada para oferecer suporte unificado a qualquer tipo de serviço que possa usufruir de endereçamento indireto.

O modelo de serviço do i3 basicamente desassocia o ato de enviar pacotes do ato de recebê-los. Para isso, cada fonte envia os pacotes para o identificador do destino, ao invés de enviar para o endereço. O destino dos pacotes demonstra interesse nesses pacotes enviando disparadores (*triggers*) contendo o seu identificador e o seu endereço. Tanto os pacotes quanto os disparadores são enviados para a estrutura de servidores do i3. Em tal estrutura, os pacotes recebidos são associados aos disparadores correspondentes para que haja o mapeamento do identificador do pacote no endereço de destino. Na estrutura i3, cada servidor é responsável por um conjunto de identificadores. Esse servidor mantém o mapeamento dos identificadores que estão sob sua responsabilidade. Logo, todo pacote recebido na estrutura é encaminhado para o servidor responsável para que ele realize o mapeamento e encaminhe o pacote para o endereço do destino correspondente. O mapeamento é mantido no servidor da estrutura i3 de maneira volátil. Assim, de tempos em tempos os destinos devem atualizar o seu mapeamento na estrutura. A Figura 2.7 ilustra todo procedimento de encaminhamento de pacotes no i3. A Figura 2.7(a) mostra o envio do disparador realizado pela estação de destino B para demonstrar interesse em um determinado conteúdo. O disparador contém o seu identificador e o seu endereço. A estrutura i3, por conseguinte, armazena o interesse. Assim, sempre que a estrutura do i3 recebe pacotes destinados ao identificador da estação B, a estrutura já sabe para que destino encaminhar, como visto na Figura 2.7(b). O endereçamento indireto proposto pelo i3 desassocia as fontes dos destinos. Assim, as fontes não precisam conhecer nem o número de destinos nem os endereços para os quais os seus pacotes são enviados, e da mesma maneira, os destinos não precisam conhecer nem o número de fontes nem os seus endereços. Uma importante característica do i3 é que um determinado identificador pode estar associado a mais de um endereço de destino. Nesse caso, o mapeamento ocorre para todos os destinos cujo identificador é igual a um número mínimo de bits do identificador mantido na estrutura de servidores do i3.



(a) Envio do disparador. (b) Encaminhamento de pacotes.

Figura 2.7. Encaminhamento de pacotes realizado pelo i3.

No caso dos nós móveis, a conexão pode ser mantida visto que ela é realizada utilizando o identificador das estações de origem e destino e não os endereços físicos. Entretanto, o nó móvel ao mudar de rede de acesso, muda de endereço físico e, portanto, deve atualizar o mapeamento na estrutura i3. Uma vez que as fontes não precisam conhecer o endereço do destino, a conexão é mantida. O i3 requer a atualização do mapeamento nos servidores, o que pode implicar problemas de latência de atualização, além de problemas de escalabilidade. Outro ponto é o aumento dos caminhos seguidos pelos pacotes em consequência do encaminhamento indireto.

O LISP (*Locator/Id Split Protocol*) [Farinacci et al. 2009b, Saucez et al. 2009] é um protocolo desenvolvido pela CISCO para lidar com os problemas dos múltiplos domicílios na Internet. O LISP é uma proposta que pode ser implementada de maneira gradual, o que representa uma vantagem. O LISP também realiza endereçamento indireto assim como o i3. Para isso, ele divide o espaço de endereçamento no espaço de endereçamento local, composto pelas redes de borda da Internet, e no espaço de endereçamento interdomínio, composto pelos roteadores da DFZ. Os roteadores dos ASes de borda da Internet possuem interfaces conectadas à DFZ configuradas com endereços denominados *Routing Locators* (RLOCs). Esses endereços são conhecidos por todos os roteadores interdomínio. Já os endereços das estações da rede local, denominados *Endpoint IDentifiers* (EIDs), possuem escopo limitado à rede de acesso. Portanto, os endereços das estações não são conhecidos pelos roteadores interdomínio e os RLOCs não podem identificar estações. O encaminhamento de pacotes fim-a-fim requer, portanto, o uso dos dois tipos de endereços. Nessa direção, os EIDs de uma rede local são associados aos RLOCs dos ISPs pelos quais possuem acesso à Internet. Um EID pode estar associado a mais de um RLOC, o que permite os múltiplos domicílios. Equivalentemente, um RLOC pode estar associado a um prefixo de endereços locais. Outra vantagem do LISP é que ele dispensa o uso de estrutura de servidores, como usado no i3. A Figura 2.8 ilustra os endereços utilizados pelo LISP. Na figura, a estação EID_A possui associado o endereço RLOC_A na DFZ. Já o EID_B possui dois endereços RLOC associados, RLOC_{B1} e RLOC_{B2}, para múltiplos domicílios.

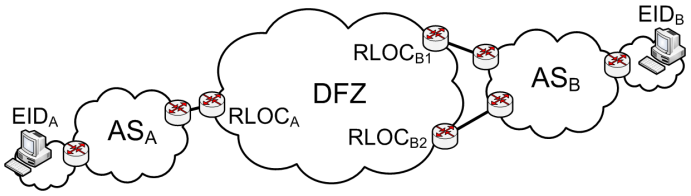
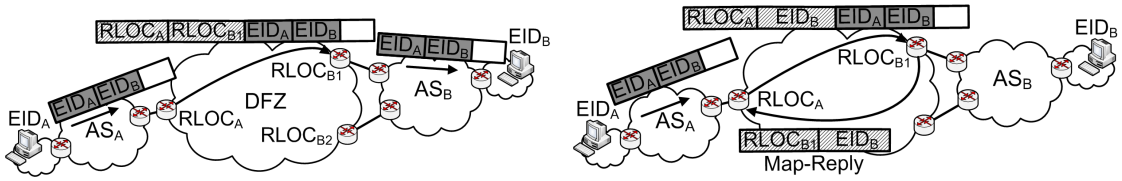


Figura 2.8. Endereçamento no LISP.

No LISP, os pacotes são encaminhados na DFZ utilizando como endereços de origem e destino os endereços dos RLOCs já que os endereços locais não são conhecidos. Entretanto, entre a estação de origem do pacote e o seu roteador de borda e entre o roteador de borda no destino e a estação de destino, os endereços usados são os EIDs de origem e destino. Para que os dois tipos de endereços sejam usados pelos pacotes, é necessário que haja o mapeamento dos EIDs nos endereços dos RLOCs. Após o mapeamento, os pacotes são encapsulados através de túneis entre os roteadores de borda. Tal procedimento

é conhecido por mapeamento e encapsulamento (*Map & Encap*). Os roteadores de borda realizam o *Map & Encap* já que eles são os únicos que conhecem a associação entre os EIDs e os RLOCs. Essa característica é uma vantagem do LISP, pois torna o seu funcionamento transparente para os outros nós da rede, além de dispensar modificações na estrutura da Internet. A Figura 2.9(a) ilustra os cabeçalhos utilizados por um pacote no LISP em cada um dos trechos percorridos: EID_A de origem até RLOC_A, RLOC_A até RLOC_{B1} e RLOC_{B1} até o destino EID_B. O mapeamento no LISP é realizado por dois sistemas, o de base de dados e o de memória cache. O de base de dados seleciona o RLOC que é usado como endereço de origem dos pacotes sendo enviados de uma rede local. Além disso, o sistema de base de dados decide se um pacote recebido da Internet deve ser desencapsulado ou não. O sistema de memória cache, por outro lado, é usado para selecionar os RLOCs das redes de acesso da estação de destino. Para tal, o mapeamento EID de destino e RLOC de destino são também armazenados nesse cache. Caso haja uma busca mal-sucedida à cache sobre o mapeamento de uma determinada estação de destino, o sistema de mapeamento do LISP faz uma requisição a um sistema de mapeamento através do protocolo de mapeamento distribuído (*Mapping Distribution Protocol*).



(a) Encaminhamento direto. (b) Uso do *Map-Reply*.

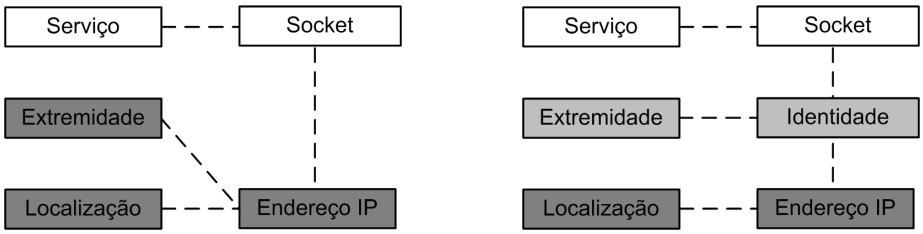
Figura 2.9. Encaminhamento de pacotes no Lisp.

No LISP, o encaminhamento de pacotes assume que a estação de origem conhece o identificador do destino e o seu próprio endereço. A resolução do identificador do destino no EID é realizada através de um sistema de resolução de nomes como o DNS. O pacote é então encaminhado pela estação de origem até o roteador de borda através de uma rota *default*. Ao chegar no roteador de borda, ele deve descobrir o mapeamento do EID do destino no RLOC correspondente. Esse procedimento depende da versão do LISP. As primeiras versões do LISP, LISP 1 e 1.5, assumem que o endereço EID é roteável na Internet. Assim, o roteador de borda da origem encapsula o pacote com um cabeçalho definido pelo LISP com o seu endereço como endereço de origem e o endereço da estação de destino como endereço de destino. Esse pacote é encaminhado pela DFZ até que chegue ao roteador de borda que conhece o mapeamento EID/RLOC do destino. O roteador de borda do destino envia uma mensagem ao roteador de borda da origem (*Map-Reply*) informando o mapeamento, como visto na Figura 2.9(b). O roteador de borda da origem armazena essa informação em sua memória cache assim que a recebe. Da mesma maneira, o roteador de borda do destino armazena o mapeamento em sua memória cache para evitar uma possível busca subsequente. As versões 2 e 3 do LISP não consideram mais que os EIDs são roteáveis na Internet. Assim, para conhecer o mapeamento EID/RLOC, o LISP utiliza um protocolo de distribuição de mapas (*Mapping Distribution Protocol*) para fazer requisições explícitas (*Map-Request*) a um serviço de mapeamento distribuído. Tais sistemas podem ser implementados baseados em sistemas de DNS, LISP versão 2; ou

em sistemas baseados em DHTs (*Distributed Hash Table*), LISP versão 3. Em ambos os casos, os sistemas respondem (*Map-Reply*) apenas às requisições recebidas.

A separação da localização da identificação proposta no LISP facilita o uso de estações multidomiciliadas visto que desassocia o endereço da estação do endereço de seu ISP. Entretanto, essa separação resulta também no emprego de sistemas de mapeamento, que introduzem um atraso para resolver EIDs em RLOCs. Como esse atraso pode ser grande, ele ainda representa um entrave à mobilidade das estações e de redes embora a separação da localização da identificação dos nós beneficie a mobilidade em um primeiro momento. O LISP possui como ponto de estudo futuro encontrar maneiras mais eficientes para lidar com a mobilidade de estações e de redes.

O HIP (*Host Identity Protocol*) [Moskowitz et al. 2008] é outro protocolo para *Loc/ID split* [OpenHIP 2009]. Para isso, ele propõe uma nova camada entre a camada de redes e a camada de transporte. Assim, a camada de transporte interage com a camada de identificação e a camada de identificação interage com a camada de rede. O HIP utiliza os conceitos de identidade e identificador. A identidade de um nó se refere à entidade abstrata que é identificada. Já o identificador se refere ao padrão binário que é utilizado no processo de identificação. Por exemplo, a identidade é o nome de uma estação (`rio.gta.ufrj.br`) e o identificador uma chave pública criptográfica. O HIP utiliza chave pública como identificador, embora qualquer tipo de identificador possa ser usado. A vantagem de utilizar uma chave pública como identificador é a possibilidade de autenticar a origem dos pacotes recebidos. Como consequência, pode-se evitar a violação dos pacotes em trânsito como feito em ataques do tipo homem-no-meio (*man-in-the-middle*). O HIP considera que identificadores que não utilizem criptografia devem ser apenas utilizados em ambientes considerados seguros. A Figura 2.10 ilustra a arquitetura atual para identificação e localização que sobrecarrega o endereçamento IP e a arquitetura proposta pelo HIP que separa a localização da identificação de uma estação.



(a) Arquietura atual da Internet. (b) Arquitetura do HIP.

Figura 2.10. Arquiteturas de identificação e localização.

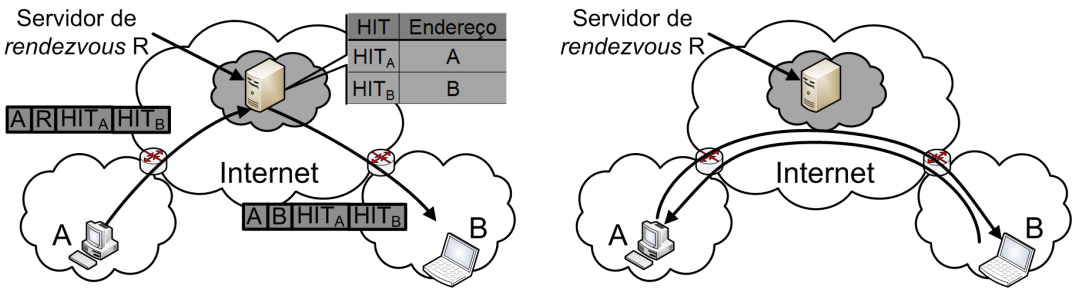
Diferente da identidade de um nó no HIP que é uma abstração, o identificador pode ser utilizado nas comunicações. Entretanto, devido ao tamanho e a variabilidade possível, a maneira escolhida pelo HIP para representar o identificador nos pacotes do protocolo é um resumo obtido através do uso de uma função *hash*. Dessa forma, a representação do identificador torna o seu uso mais simples já que possui tamanho fixo, além de significar um formato consistente independente dos algoritmos criptográficos empregados. O resultado da função *hash* é denominado como rótulo da estação (*Host Identity Tag* - HIT) e, consequentemente, deve ser único globalmente.

Uma comunicação com HIP é iniciada a partir de um processo de autenticação e troca de chaves. Portanto, sempre antes de enviar dados é realizada uma troca de base através do BEX (*Base Exchange Protocol*) e uma associação HIP se torna ativa. Esse protocolo é executado fim-a-fim, ou seja, entre as estações finais. O procedimento de autenticação envolve a troca de desafios criptográficos para autenticação. Todo o procedimento pode ser atualizado quando as chaves criptográficas expiram ou quando uma estação se desloca.

A camada de transporte estabelece conexões utilizando os HITs de origem e destino que são oferecidos pela camada introduzida pelo HIP para identificação das estações envolvidas na comunicação. Entretanto, para os pacotes serem encaminhados na Internet, a camada introduzida pelo HIP precisa substituir os HITs pelos endereços IP correspondentes. Para isso, a estação de origem dos pacotes substitui o HIT de origem pelo endereço IP correspondente e obtém via um sistema de resolução de nomes, o endereço IP do destino. No HIP, a resolução de nomes não retorna apenas o endereço IP da estação, mas também informações como o HIT e a chave pública. Após a resolução de nomes, a comunicação de dados inicia-se entre a estação de origem e destino. Uma vez descoberto o endereço de um nó, esse endereço é utilizado até que a comunicação se encerre ou que o nó de destino se desloque. Essa característica difere o HIP do LISP, já que no LISP o encaminhamento dos pacotes sempre utiliza túneis. No HIP, se uma estação se mover, a própria estação pode avisar à estação de destino qual o seu novo endereço físico. Entretanto, se uma nova estação desejar estabelecer uma conexão com a estação que se moveu, a resolução do nome irá retornar o endereço IP antigo já que o mapeamento no DNS não foi atualizado. Para esses casos, o HIP usa servidores na Internet para evitar que as estações móveis precisem sempre atualizar o DNS.

O HIP define um servidor estático chamado de *rendezvous* para auxiliar em um processo de encaminhamento indireto, semelhante ao usado no i3. Para isso, o nome da estação móvel é mapeado pelo DNS no endereço IP do *rendezvous*. O servidor *rendezvous* possui atualizada a localização da estação móvel. Parte-se da premissa que é melhor atualizar o endereço IP em um servidor na Internet que o DNS. Logo, uma estação da Internet que deseja se comunicar com uma estação móvel realiza a resolução de nomes e recebe o endereço IP do servidor de *rendezvous* e o HIT da estação móvel. A estação da Internet inicia o processo de autenticação e troca de chaves utilizando o endereço IP e o HIT recebido. A mensagem é encaminhada até o servidor de *rendezvous* que ao perceber que o HIT da mensagem não é um dos seus, verifica em seus registros se possui o mapeamento correspondente. Todas as estações que desejam utilizar um servidor de *rendezvous* devem se registrar nesses servidores. Portanto, caso a mensagem seja destinada a um HIT cujo mapeamento no endereço IP é conhecido, o servidor encaminha o pacote até a estação móvel correspondente. Ao receber a primeira mensagem de autenticação e troca de chaves, a estação móvel responde diretamente à estação de origem e o restante da comunicação passa a ser direta desde então sem o intermédio do *rendezvous*. Se a estação móvel mudar de rede novamente, ela atualiza a outra estação e o servidor *rendezvous* da sua nova localização. A Figura 2.11 ilustra o funcionamento do HIP com o uso do servidor de *rendezvous*. Na Figura 2.11(a), a estação de origem A envia a primeira mensagem com o endereço IP do servidor de *rendezvous* R obtido após a resolução do nome no DNS. O servidor de *rendezvous*, então, encaminha a mensagem até a estação

móvel. Em seguida, a estação móvel se comunica diretamente com a estação que originou a comunicação, como visto na Figura 2.11(b).



(a) Encaminhamento através do *rendezvous*. (b) Encaminhamento direto.

Figura 2.11. Encaminhamento de pacotes no HIP.

Apesar do HIP encaminhar pacotes de uma maneira diferente do LISP, ele também pode sofrer problemas com a mobilidade de estações devido à latência de atualização do mapeamento. Em redes móveis em velocidades altas, essa latência pode não ser suficiente para atender as demandas de serviços com restrições de atraso. O HIP provê ainda suporte a múltiplos domicílios, já que mais de um endereço pode estar relacionado à mesma identidade [Nikander et al. 2008].

O Shim6 [Nordmark e Bagnulo 2009] é uma proposta para proporcionar tolerância a falhas a estações multidomiciliadas em domínios IPv6. O Shim6 mantém comunicações correntes mesmo em caso de falha de alguns localizadores. Caso uma parte das conexões de uma rede, as outras ativas devem manter a comunicação. No pior caso, se todas as conexões falharem, o tempo de restabelecimento da comunicação é igual ao tempo de atualização do DNS e das atualizações se propagarem pela rede. Nesse sentido, o Shim6 possibilita o uso dos múltiplos domicílios. Semelhante ao HIP, o Shim6 atribui um nome especial para o identificador do nó, chamado de ULID (*Upper Layer IDentifier*). Já o localizador da estação não possui nome especial e se refere ao endereço IPv6 da camada de rede. Diferente das propostas anteriores, entretanto, o endereço IPv6 utilizado no primeiro contanto com a outra estação cumpre também o papel de ULID. Logo, o Shim6 não propõe o uso de um nome espaço de identificação. Porém, caso o localizador mude ao longo da comunicação entre os nós, o ULID permanece constante. Essa possibilidade é considerada pouco frequente visto que o Shim6 não foi desenvolvido diretamente para redes cujos nós mudem de localização dinamicamente, como seria o caso das redes móveis.

O Shim6 é representado como uma camada diretamente acima da camada de rede. Assim, tanto as aplicações quanto os protocolos de camadas acima da camada Shim6 utilizam os ULIDs que foram mapeados a partir dos localizadores das estações de origem e destino. A camada Shim6 mantém o mapeamento por par de ULIDs. Portanto, todas as conexões envolvendo o mesmo par origem-destino compartilham o mesmo mapeamento. O mapeamento deve ser realizado de maneira consistente na origem e no destino para que os protocolos de camadas superiores vejam os pacotes sendo enviados utilizando os ULIDs fim-a-fim. Essa característica é mantida mesmo que os pacotes sejam encaminhados pela rede utilizando os localizadores como endereços IP e mesmo que os localiza-

dores mudem. Além disso, o mapeamento realizado para o par de nós deve ser mantido independentemente do protocolo de camada superior e de qualquer outra conexão estabelecida. Os protocolos de camadas superiores que estão cientes da operação do Shim6 podem associar mais de um par de localizadores a um único par de ULIDs.

No Shim6, uma determinada estação de origem inicia uma transmissão caso alguma de suas aplicações possua pacotes a enviar. Nesse momento, o Shim6 pode se comunicar com o destino do pacote para conhecer possíveis pares alternativos de localizadores associados ao par de ULIDs. Essa comunicação é possível visto que o ULID é igual ao localizador da estação, o que dispensa o uso de sistemas de mapeamentos. Os múltiplos pares de localizadores são utilizados para casos de falha. Se uma comunicação falhar, o Shim6 pode testar novos pares de localizadores até que um deles restabeleça a comunicação. Ao encontrar esse par, os localizadores são utilizados nos próximos pacotes, mas uma extensão de cabeçalho é utilizada para que o destino saiba qual o ULID correspondente ao novo par de localizadores. Essa extensão é chamada de rótulo de contexto (*Context Tag*). O uso do rótulo de contexto permite que as mudanças dos localizadores permaneçam transparentes para os protocolos de camadas superiores.

O fato de o localizador inicial ser igual ao identificador pode levar a períodos de instabilidade. Caso o localizador de uma estação mude, nada impede que o localizador anterior seja atribuído a outro nó da rede. Assim, há a possibilidade de duas estações diferentes utilizarem o mesmo identificador já que o identificador é inicialmente igual ao localizador. O Shim6 evita essa possibilidade forçando que as comunicações sejam terminadas quando uma ULID se tornar inválida. Um mecanismo de recuperação de contexto é então acionado para que a outra estação na comunicação fique ciente que a ULID não está mais relacionada com o localizador anterior. Esse procedimento é uma consequência de se evitar sistemas de mapeamento, como p. ex. o *rendezvous* do HIP.

O mapeamento entre identificadores e localizadores ou entre endereços locais e endereços intradomínio é um problema em aberto para os diferentes protocolos de *Loc/ID Split*. No LISP, por exemplo, muitos sistemas são avaliados para tornarem o sistema de mapeamento escalável e seguro. Os sistemas de mapeamento podem ser classificados em PUSH ou PULL. Nos sistemas PUSH, o mapeamento mantido por elementos centralizadores é enviado aos roteadores sem que os roteadores façam requisições. Essa estratégia pode aumentar a carga de controle enviada na rede, além de manter todos os mapeamentos conhecidos pelos roteadores de borda. O ponto positivo é que o mapeamento não sofre atrasos. Propostas como o NERD (*Not-so-novel EID to RLOC Database*) [Lear 2010] anunciam os mapeamentos conhecidos a todos os roteadores da rede sem a necessidade de requisições. Os sistemas PULL funcionam através de requisições. Portanto, sempre que um roteador de borda não conhecer um determinado mapeamento, ele envia uma requisição para o sistema de mapeamento. Tais sistemas de mapeamento podem funcionar como um DNS ou uma DHT (*Distributed Hash Table*) [Iannone e Bonaventure 2007]. Iannone e Bonaventure demonstram que sistemas de mapeamento baseados em DNS não sofrem com problemas de escala [Iannone e Bonaventure 2007]. Para isso, os autores demonstram que é possível controlar o tamanho da memória cache dos nós a partir do ajuste do tempo de permanência dos mapeamentos na memória. Luo *et al.* [Luo et al. 2009] demonstram que sistemas baseados em DHTs também podem alcançar desempenho satisfatório em termos de escalabilidade. Sistemas como o LISP-ALT [Farinacci et al. 2009a]

e LISP-CONS [Brim et al. 2008] são sistemas híbridos, ou seja, que utilizam tanto a estratégia PULL quanto a PUSH. O LISP-ALT define uma arquitetura na qual alguns roteadores são responsáveis por manter os mapeamentos. Esses roteadores trocam informações baseado na estratégia PUSH. Já os roteadores de borda requisitam o mapeamento aos roteadores responsáveis. O LISP-CONS difere do LISP-ALT pela organização dos roteadores responsáveis por manter os mapeamentos. No LISP-CONS, a estrutura desses roteadores é hierárquica. Portanto, os mapeamentos são trocados entre os roteadores em nível hierarquicamente superior através da estratégia PUSH. Já os roteadores de borda requisitam os mapeamentos aos roteadores de nível hierárquico inferior que consultam os roteadores em nível superior também através de requisições.

2.4.2. Roteamento plano

O roteamento plano é outra possibilidade para contornar os problemas oriundos da semântica sobrecarregada do endereço IP. Para isso, o roteamento plano extrapola o conceito da separação entre localização e identificação removendo o conceito de localização. O objetivo é rotear pacotes baseado apenas na identificação dos nós. Assim, torna-se necessário garantir a unicidade do identificador e não mais a relação entre a localização do nó e a topologia de rede. Uma vez que o roteamento seja baseado no identificador, e não mais em um endereço correlacionado com a topologia de rede, esse nó pode se deslocar sem prejuízo das suas conexões. Além disso, não realizar a resolução de nomes evita o emprego de infraestrutura de rede, p. ex. DNS, e aumenta a disponibilidade da comunicação já que o sistema de resolução de nomes não é mais um ponto de falha. O roteamento plano deve ser investigado apesar da premissa da Internet de manter no cabeçalho dos pacotes informação estruturada da localização das estações. Entretanto, trabalhos da área demonstram que esse tipo de proposta ainda enfrenta problemas de escalabilidade.

Até o momento, o ROFL (*Routing On Flat Labels*) [Caesar et al. 2006b] é o principal trabalho em roteamento plano. O ROFL utiliza rótulos para identificação dos nós e desassocia a identificação da localização na rede. A operação do ROFL é baseada no conceito de DHTs (*Distributed Hash Tables*), utilizadas comumente por protocolos de redes par-a-par. A motivação principal do uso das DHTs é a sua capacidade de distribuição homogênea de carga ou funcionalidades. Em uma rede par-a-par, p. ex., é desejável que os dados armazenados pelos nós da rede sejam distribuídos homogeneamente para evitar nós sobrecarregados. A característica de distribuição homogênea, por conseguinte, leva as DHTs a serem aplicadas em sistemas que evitam estruturas hierárquicas. Por isso, o ROFL utiliza as DHTs no roteamento ao invés de distribuição de conteúdo.

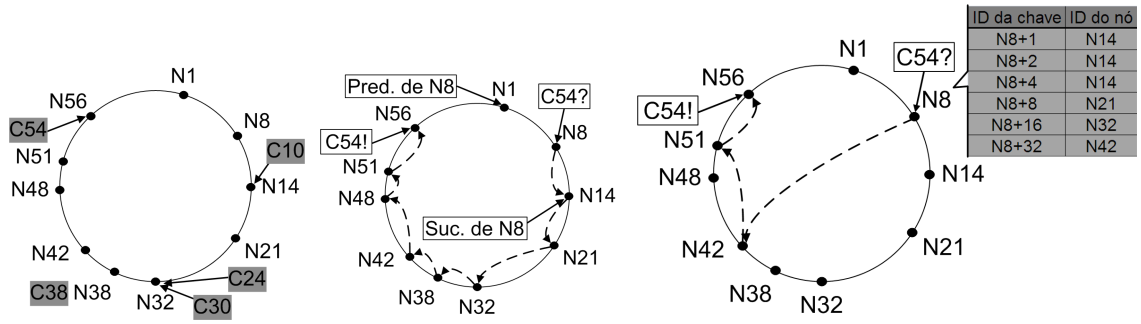
As propostas para roteamento plano, incluindo o ROFL, devem funcionar tanto para comunicações entre nós no mesmo AS quanto para nós em ASes diferentes. Portanto, as propostas devem definir protocolos para o estabelecimento e a manutenção das comunicações em ambos os casos. Os protocolos propostos pelo ROFL são baseados no Chord [Stoica et al. 2003] e no Canon [Ganesan et al. 2004], que são sistemas de distribuição que utilizam DHTs. O primeiro é utilizado pelo ROFL como base para as comunicações intradomínio. Já o segundo, é utilizado como base para as comunicações interdomínio. Antes de apresentar o funcionamento do ROFL é interessante apresentar o funcionamento do Chord e do Canon.

O Chord é um protocolo originalmente proposto para busca de conteúdo em redes par-a-par. No Chord, cada tipo de dado é associado a uma chave que é escolhida em função do identificador utilizado na camada de aplicação, p. ex. um nome de um arquivo em um sistema par-a-par. Já na camada de rede, o Chord associa um identificador a cada uma das chaves e também associa identificadores a cada um dos nós. Ambos os identificadores são obtidos com o uso de funções *hash*. Por fim, o Chord mapeia um identificador de chave a um nó da rede. Assim, cada nó fica responsável por uma ou mais chaves. Caso a chave identifique um dado, este pode ficar armazenado no nó correspondente. As principais vantagens do Chord são a distribuição uniforme de chaves que evita a sobrecarga de algum nó em particular e o armazenamento de informações locais de roteamento em cada nó. O objetivo é aumentar a escalabilidade do sistema ao dividir a carga de armazenamento de dados e de informação de controle.

O mapeamento entre os identificadores das chaves e dos nós é realizado em um espaço de identificação circular de módulo 2^m, onde m é o tamanho do identificador em bits. Assim, um identificador de chave `C` é mapeado no primeiro nó cujo identificador é igual a `C` ou, se o nó com identificador igual não existir, ao primeiro nó encontrado seguindo o espaço de identificação no sentido horário. A Figura 2.12(a) ilustra um espaço de identificação de módulo 2^6. Na figura, os identificadores que começam com `C` são os identificadores das chaves e os que começam com `N`, os dos nós. A figura mostra que o identificador da chave `C10` não foi diretamente mapeado em um nó com um identificador de mesmo número. Logo, `C10` foi mapeada no próximo nó no sentido horário, ou seja, no nó `N14`. Já o nó `N38` possui um identificador de chave que coincide com o seu próprio identificador. O processo de localização básico de um identificador de chave consiste em uma busca no espaço circular. Essa busca é realizada salto-a-salto, no qual cada salto é um identificador de nó. A localização termina quando o identificador da chave pertencer ao intervalo entre o identificador do nó analisado e o seu próximo nó no sentido horário. Nesse caso, a busca retorna o identificador desse próximo nó. No Chord, cada nó armazena o identificador do próximo nó no sentido horário e do anterior. Esses nós são chamados de nó sucessor e predecessor, respectivamente. A Figura 2.12(b) mostra como é realizado o processo de localização de um identificador de chave no Chord a partir do nó `N8`. Na figura, os nós `N1` e `N14` são os nós sucessor e predecessor do nó `N8`, respectivamente. O nó `N8` busca a localização do identificador da chave `C54`. O processo de localização é repetido pelos nós no círculo até encontrar a chave desejada.

O espaço de identificação circular permite que o nó responsável por um determinado dado seja sempre encontrado, já que a busca pode apenas ser realizada no sentido horário. Na Internet, entretanto, a rede é organizada hierarquicamente e a busca por um determinado nó pode exigir mais estados por roteador. Enquanto no Chord cada nó precisa apenas conhecer o seu sucessor e predecessor no espaço circular, na Internet, cada nó precisa saber o próximo salto para cada possível destino. A vantagem da redução de estados proposta pelo Chord, entretanto, insere um problema. Quanto menos estados forem armazenados por nó, maior é a quantidade de saltos que o procedimento de localização deve realizar. Esse compromisso é abordado pelo Chord ao introduzir uma tabela de roteamento por nó. No Chord, cada nó mantém uma tabela de m linhas para agilizar o processo de localização dos identificadores das chaves. Cada linha da tabela identifica o nó responsável pelos identificadores de chaves distantes até 2^m. Portanto, cada linha

da tabela é uma tupla composta pelo identificador da chave (igual ao número do identificador do nó realizando a busca mais 2^{i-1}, onde $1 \leq i \leq m$) e pelo identificador do nó responsável por essa chave. Caso o identificador da chave seja maior que o último identificador na lista, a busca é adiantada até o primeiro nó no sentido horário após a lista. A Figura 2.12(c) ilustra o procedimento de localização de `C54` realizado por `N8` com o uso de tabela de roteamento. O nó `8` adianta a sua busca até `N42` já que o identificador da chave buscada é maior que o identificador do último nó da tabela. O nó `42` também adianta a localização saltando `N48`. Ao combinar o espaço de identificação circular com a tabela de roteamento, o Chord propõe uma solução para localização de identificadores de nós que aumenta a escalabilidade do sistema. A ideia de organizar a rede em um espaço de identificação circular pode ser aplicada à Internet tomando-se o cuidado para que o requisito de unicidade de identificação seja garantido. Uma vez que um nó tenha um identificador único associado, independente da sua localização, seu identificador permanece o mesmo e a sua posição no espaço de identificação também. Isso permite que os nós possam se deslocar sem quebrar as conexões estabelecidas.

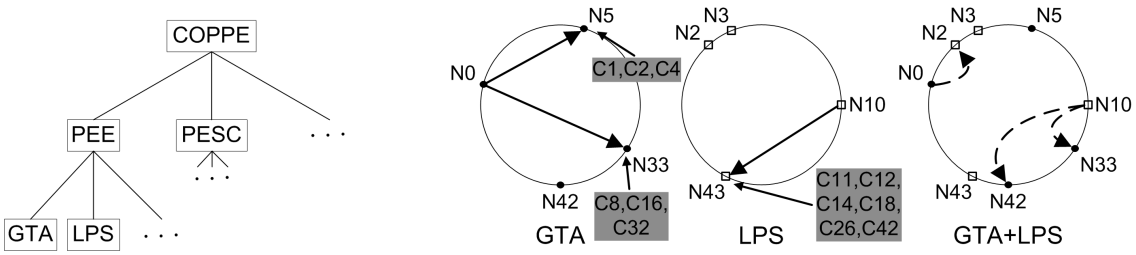


(a) Mapeamento entre identificadores de chave e nó. (b) Busca de identificador no Chord. (c) Busca de identificador no Chord usando tabelas.

Figura 2.12. Funcionamento do Chord.

Um problema do uso das DHTs como no Chord é que, embora o seu uso possibilite a distribuição balanceada de carga entre os nós, o projeto de sistemas escaláveis está normalmente associado a sistemas com estruturas hierárquicas. Portanto, uma maneira de aumentar a escalabilidade do sistema é aplicar hierarquia sem perder as principais vantagens das DHTs. O Canon [Ganesan et al. 2004] é uma proposta para estender o Chord adicionando hierarquia à sua estrutura. Embora contraditório, o objetivo é construir um sistema híbrido para reunir as principais vantagens das DHTs e dos sistemas com estruturas hierárquicas. Uma das motivações é a adaptação mais suave à realidade da Internet que é organizada de maneira hierárquica e dividida em diferentes ASes. A Figura 2.13(a) mostra um exemplo ilustrativo da estrutura física da rede da COPPE, que é a instituição responsável pela pós-graduação em engenharia da UFRJ. A COPPE é dividida em diferentes programas (Engenharia Elétrica, Mecânica etc.) e cada programa possui laboratórios afiliados. Portanto, a organização das redes segue a mesma estrutura hierárquica como na Internet, onde cada uma das redes representa um domínio. A estrutura utilizada é uma árvore na qual os laboratórios GTA e LPS são folhas e a raiz é a COPPE.

No Canon, cada domínio folha na árvore define um espaço de identificação circular, assim como no Chord. A diferença está nos domínios internos da árvore que definem



(a) Hierarquia entre domínios. (b) União de espaços de endereçamento.

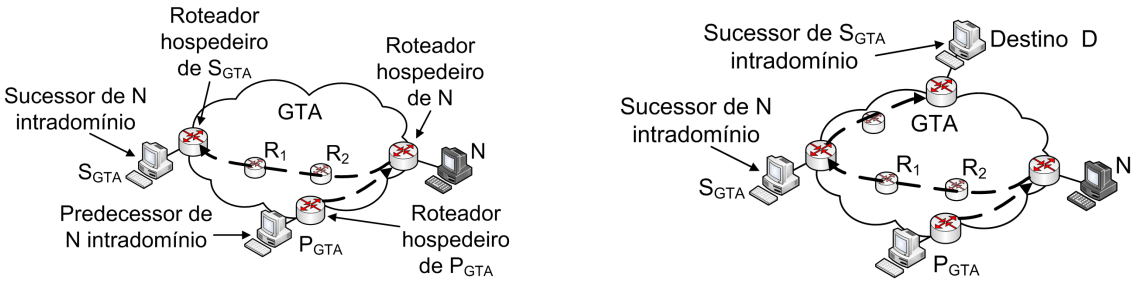
Figura 2.13. Funcionamento do Canon.

um espaço de identificação circular que une os seus nós e mais os nós dos seus domínios filhos. Essa estrutura é repetida hierarquicamente até alcançar o domínio raiz, no exemplo da Figura 2.13(a), o domínio da COPPE. A identificação dos nós deve ser única globalmente e cada nó deve criar uma tabela de roteamento baseada no módulo do espaço de identificação circular da rede. A Figura 2.13(b) ilustra o espaço de identificação circular módulo 2^6. O nó `N0` do GTA possui entradas em sua tabela de roteamento para identificadores de chaves que estão associadas aos nós `N5` (chaves 1, 2 e 4) e `N33` (chaves 8, 16 e 32). Já o nó `N10` do LPS possui entradas para identificadores de chaves associadas ao nó `N43` (chaves 11, 12, 14, 18, 26 e 42). A união do espaço de endereçamento do GTA com o do LPS mantém todas as entradas estabelecidas. Com a união, novas entradas são adicionadas seguindo duas regras pré-definidas. A primeira regra define que um nó `N` do GTA pode criar uma ligação com um nó `N'` do LPS se e somente se `N'` for o nó mais próximo que esteja distante de no máximo 2^m. Já a segunda regra define que o nó `N'` que atende o requisito da regra anterior deve ser mais próximo de `N` do que qualquer outro vizinho de `N` do GTA. Essa segunda condição limita o número de entradas aos nós do espaço de endereçamento vizinho mais próximos dos identificadores de chaves entre $\mathtt{N} < \mathtt{N} + 2^{i-1} < \mathtt{N} + 2^m$ do que os nós do próprio espaço. Na Figura 2.13(b), os espaços de endereçamento unidos demonstram as entradas da tabela de roteamento que foram formadas pelos nós `N0` e `N10` com o processo de união de espaços. Note que `N0` pode manter entradas para `N2` devido às distâncias 1 e 2 e para `N10` devido à distância 8, de acordo com a primeira regra. Entretanto, `N10` está mais distante de `N0` que `N5` que pertence também ao GTA. Portanto, a entrada correspondente não é adicionada por `N0`, como definido pela segunda regra.

A união de espaços circulares é feita recursivamente desde os nós folhas até o nó raiz. Ao final, o espaço de endereçamento circular é formado englobando todos os nós da rede. O roteamento realizado no Canon segue uma abordagem gulosa na qual um determinado pacote é encaminhado até o identificador no mesmo nível hierárquico mais próximo ao destino. Por exemplo, na Figura 2.13(b), os pacotes enviados pelo nó `N2` ao nó `N42` seguem o caminho no LPS até `N10` e de `N10` até `N42`. É importante mencionar que as entradas existentes na tabela de roteamento de cada nó são preservadas mesmo após o processo de união. Além disso, o processo realizado pelo Canon pode levar em conta domínios que pertençam a ASes diferentes.

Semelhante ao Chord, cada nó no ROFL possui um nó sucessor e um nó predecessor na rede intradomínio. Entretanto, o ROFL considera que os nós pertencem a

uma rede local de acesso que oferece conectividade via um roteador de interconexão, denominado pelo ROFL como roteador hospedeiro. Esse roteador armazena os caminhos completos que o ligam até os roteadores hospedeiros dos nós sucessores e predecessores dos seus nós hóspedes. Da mesma maneira, os roteadores hospedeiros dos nós sucessores e predecessores também armazenam o caminho completo no sentido reverso. O caminho completo é definido como a sequência salto-a-salto dos identificadores de todos os roteadores físicos entre hospedeiros e é utilizado durante o encaminhamento de pacotes para roteamento pela fonte. A Figura 2.14(a) ilustra os nós sucessor ($\mathtt{S}_{GTA}$) e predecessor ($\mathtt{P}_{GTA}$) de $\mathtt{N}$ no domínio do GTA. Note que tanto o caminho armazenado até o nó sucessor (N, R_1, R_2, S_{GTA}) quanto o caminho até o nó predecessor (N, P_{GTA}) iniciam e terminam nos roteadores hospedeiros. A rede também pode possuir roteadores que originam ou recebem tráfego. Nesse caso, os roteadores também possuem sucessores e predecessores e os caminhos são armazenados semelhantemente ao caso dos nós hóspedes. O roteador hospedeiro é responsável por interconectar um nó que entra na rede ao espaço de identificação circular. Para tanto, o roteador hospedeiro primeiramente autentica o nó entrante, para somente após encontrar os roteadores sucessores e predecessores desse nó no espaço circular. Caso o roteador hospedeiro falhe, os nós hóspedes devem perceber isso através de temporizadores expirados. Nesse caso, os nós hóspedes procuram um roteador alternativo para entrar novamente na rede ou elegem um roteador hospedeiro de reserva. Essa última opção só é possível se mais de um roteador hospedeiro for escolhido desde a primeira entrada do nó hóspede.



(a) Nós sucessores e predecessores.

(b) Encaminhamento de pacotes.

Figura 2.14. Funcionamento do ROFL intradomínio.

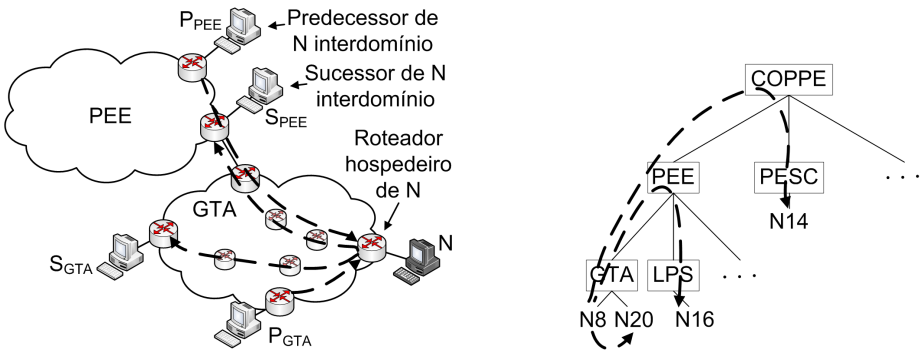
Os roteadores hospedeiros podem usar memória cache para armazenar ponteiros para identificadores de nós. Cada ponteiro é uma referência aos caminhos para os nós sucessores e predecessores dos seus nós hóspedes. Os roteadores hospedeiros também podem armazenar ponteiros referentes a caminhos que passam através dele vindos de outros nós da rede que não sejam nem sucessores ou predecessores dos seus hóspedes. O primeiro tipo de caminho possui precedência sobre o posterior, uma vez que a cache é limitada em tamanho. Assim como no Chord, o roteamento é realizado de maneira gulosa auxiliada por uma tabela de roteamento que possui ponteiros para identificadores de nós. Logo, um pacote enviado para um determinado identificador é encaminhado através do ponteiro mais próximo daquele identificador de destino. No pior caso, um pacote pode ser encaminhado ao longo de todos os sucessores. O conceito de proximidade entre ponteiros está relacionado com a proximidade entre identificadores. Os roteadores hospedeiros mantêm uma lista ordenada de identificadores e sempre que vão encaminhar

um pacote, eles comparam o identificador com a sua lista. O roteador que possuir o identificador menor mais próximo do identificador do destino é utilizado como próximo salto. A Figura 2.14(b) ilustra o encaminhamento de pacotes salto-a-salto. O pacote é encaminhado entre a origem N e o destino D passando por todos os nós sucessores dos roteadores hospedeiros no caminho.

O ROFL define procedimentos para casos de falhas de roteadores. Em caso de falha, os roteadores vizinhos inspecionam todos os seus ponteiros em cache e enviam mensagens de falha ao longo de qualquer caminho que possua o roteador com problemas. Se for um problema de um nó hóspede, o roteador envia mensagens de falha aos predecessores e sucessores do nó que falhou. Quando uma mensagem de falha chega a roteadores hospedeiros, eles recuperam os nós sucessores e predecessores do nó com problemas na tentativa de manter a rede conectada. Entretanto, nem sempre isso é possível. Por isso, os roteadores sempre distribuem rotas aos roteadores mais estáveis para evitar particionamentos da rede. Os roteadores localmente desempenham checagens de correção baseados no conteúdo das mensagens recebidas e então executam protocolos de recuperação de particionamentos para assegurar que a rede convirja em um único espaço de endereçamento.

No interdomínio, a operação do ROFL é semelhante à do intradomínio, exceto por considerar políticas no nível de ASes. Tais políticas podem ser modeladas como um grafo de ASes organizado de maneira hierárquica. Cada AS forma um espaço de identificação circular e os diferentes ASes se comunicam através da estrutura em árvore interdomínio. Para possibilitar a comunicação entre os diferentes ASes, os espaços de identificação devem se unir. O processo de união entre ASes diferentes pode ser dividido em três etapas. Na primeira etapa, cada AS deve descobrir todos os outros ASes hierarquicamente superiores com os quais possui acordos. Na segunda etapa, os ASes se unem recursivamente, assim como no Canon, utilizando os roteadores que compartilham enlaces em comum em ASes diferentes. Na última etapa, cada AS define ponteiros para roteadores em ASes vizinhos que diminuam o caminho entre eles. Um determinado nó em um AS pode ser globalmente alcançável se o seu roteador hospedeiro mantiver predecessores e sucessores referentes em cada nível de hierarquia de ASes, como ilustrado na Figura 2.15(a). Semelhante ao caso intradomínio, no nível dos ASes o roteamento também é realizado pela fonte. Assim, um determinado pacote é encaminhado de acordo com o caminho de ASes pré-estabelecido armazenado nos roteadores hospedeiros como um vetor de caminho do BGP. Caso haja falha nos enlaces entre ASes, os ASes restabelecem a sub-árvore formada entre eles para garantir que a rede mantenha conectividade. Se isso não for possível, enlaces operacionais podem ser adicionados. As políticas entre diferentes sistemas autônomos permitem que alguns caminhos sejam tratados como caminhos de reserva.

O encaminhamento de pacotes no ROFL segue propriedades de isolamento. Uma comunicação entre nós no mesmo AS não utiliza ponteiros para nós externos àquele AS. Semelhantemente, uma comunicação entre nós em ASes diferentes resulta em pacotes encaminhados através do menor caminho possível na árvore. Logo, o caminho formado entre dois ASes deve seguir através do primeiro ascendente comum. A Figura 2.15(b) ilustra o isolamento entre os ASes no ROFL. Caso `N8` queira se comunicar com `N20`, ele o fará sem utilizar ponteiros externos. Entretanto, caso `N8` queira se comunicar com `N16`, `N8` precisa utilizar o nó sucessor e predecessor no domínio do PEE, subindo até o primeiro nível hierárquico em comum. Equivalentemente, a comunicação entre `N8` e `N14`



(a) Nós sucessores no ROFL.

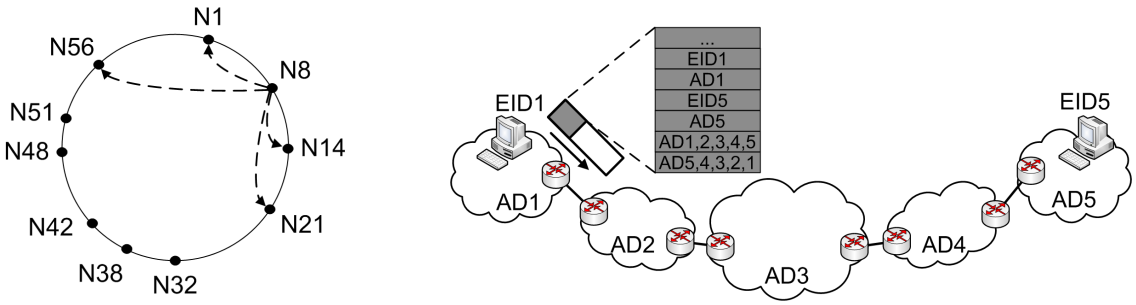
(b) Hierarquia de encaminhamento no ROFL.

Figura 2.15. Funcionamento do ROFL interdomínio.

utiliza o domínio da COPPE para a comunicação. Como visto, o encaminhamento de pacotes segue uma estrutura hierárquica. Isso demonstra que apesar de utilizar um espaço de identificação plano, questões como escalabilidade e divisão da Internet em ASes ainda requerem que o roteamento seja feito de maneira estruturada hierarquicamente.

Os problemas de escalabilidade do ROFL estão ligados à quantidade de ponteiros armazenados para outros nós da rede, à carga de controle para o estabelecimento dos espaços de identificação circulares, à latência da entrada dos nós e recuperação de falhas. Além do ROFL, outro exemplo de protocolo que planifica o espaço de identificação é o VRR (*Virtual Ring Routing*) [Caesar et al. 2006a]. O VRR, porém, limita o escopo do roteamento a uma rede sem-fio de múltiplos saltos. Assim, o VRR reduz o problema da escalabilidade já que lida com um número menor de nós se comparado à Internet. O funcionamento do VRR é semelhante ao do ROFL intradomínio. No VRR, entretanto, cada nó N mantém uma tabela de caminhos para os m vizinhos consecutivos mais próximos no espaço de identificação circular. Desses m vizinhos, metade possui identificadores no sentido horário e a outra metade no sentido anti-horário ao de N. A Figura 2.16(a) mostra os m vizinhos de `N8`, $m = 4$. Cada caminho é definido pelos identificadores de origem e destino e pelos identificadores do próximo salto na rede física. A tabela de caminhos do VRR é diferente da tabela de roteamento do ROFL pois armazena nós sucessores e predecessores consecutivos. Além disso, o VRR utiliza informações da rede física já que foi desenvolvido para uso em redes sem-fio, diferente do ROFL. No VRR, essa informação evita que vizinhos físicos cujos enlaces não ofereçam qualidade mínima sejam usados como próximos saltos.

Diferente do ROFL e do VRR, mas ainda na direção da planificação do espaço de endereçamento, está o AIP (*Accountable Internet Protocol*) [Andersen et al. 2008]. O AIP é outra proposta para roteamento na Internet que evita o uso de prefixos e endereços sem classes definidas (*Classless Inter-Domain Routing* - CIDR). Para isso, ele retorna à estrutura original de endereços da Internet que era a concatenação do identificador de rede com o da estação. Essa estrutura possuía dois níveis hierárquicos, um de rede, utilizado pelos roteadores para encaminhar pacotes, e outro de estações. Logo, O AIP utiliza uma abordagem com dois níveis hierárquicos, diferente do ROFL e do VRR. No AIP, o nível dos roteadores é chamado de domínio de responsabilização (*Accountability Domain* - AD) e o das estações é chamado de identificador de pontos finais (*End-point IDentifier*



(a) Vizinhos no espaço de identificação circular no VRR.

(b) Encaminhamento no AIP.

Figura 2.16. Outros protocolos de roteamento plano.

- EID). Essa divisão reduz os problemas de escalabilidade comparado ao ROFL, mas não desassocia completamente o identificador do nó do identificador da rede. Um dos problemas enfrentados pelo AIP é como relacionar os ASes aos domínios de responsabilização. Na Internet, os ASes possuem identificadores totalmente desassociados dos prefixos de endereços anunciados por eles. No AIP, o identificador do domínio de responsabilização é utilizado nos dois casos.

A estrutura de endereçamento do AIP é formada pela concatenação do identificador do domínio (AD) com o identificador do nó (EID). O AD e o EID são formados a partir de saídas de uma função *hash*. Ambos são calculados sobre a chave pública que identificam o domínio e a estação, respectivamente. Essa característica permite que os endereços possam ser autenticados. As conexões TCP são estabelecidas utilizando apenas o identificador do nó, assim há a possibilidade dos nós se deslocarem entre domínios diferentes. O único requisito é a garantia da unicidade do identificador do nó, conquistada utilizando parte do endereço MAC no identificador. O encaminhamento de pacotes no AIP é feito utilizando roteamento pela fonte no nível dos domínios. Isso ocorre porque tanto a ausência de organização hierárquica dos domínios quanto o não emprego de prefixos de rede impedem que o encaminhamento seja feito por técnicas de busca de melhor prefixo (*best match prefix*). A Figura 2.16(b) mostra o caminho percorrido por um pacote desde sua origem EID1 até o seu destino EID5. O cabeçalho do pacote possui entre outros campos, os identificadores do nó e do domínio de origem, os identificadores do nó e do domínio do destino e as pilhas de identificadores de domínios entre a origem e o destino tanto no caminho direto quanto no reverso.

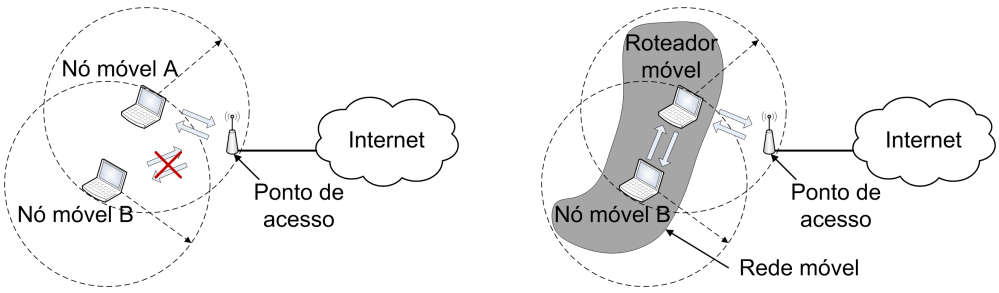
O AIP, assim como o ROFL, ainda possui problemas de escalabilidade e aplicabilidade na Internet. Em ambos os casos, a infraestrutura da Internet desde sistemas finais até roteadores precisariam ser alterados para adotar qualquer uma das propostas. Portanto, o emprego dessa estratégia ainda não é óbvio apesar de solucionar problemas como a mobilidade de estações.

2.4.3. Mobilidade de Rede

A mobilidade está normalmente associada à mobilidade de uma única estação e não à mobilidade de uma rede, onde uma rede é um conjunto de uma ou mais estações conectadas. Um exemplo disso é o modelo base do IP móvel que define a existência de um agente estrangeiro, um agente domiciliar, uma estação correspondente e uma estação

móvel. Esse modelo representa uma limitação já que a popularização das redes sem-fio requer que o suporte à mobilidade seja estendido aos casos nos quais redes inteiras se movem e não apenas uma única estação. Exemplos típicos de redes móveis são as redes ad hoc, as redes veiculares [Alves et al. 2009] e até mesmo as redes formadas entre dispositivos de comunicação baseados no IP carregados por pessoas (*Personal Area Network* - PAN). Em todos esses cenários as diferentes estações podem se deslocar em conjunto e podem estar conectadas entre si com acesso à Internet. É válido ressaltar que o mais importante é manter as conexões previamente estabelecidas ativas mesmo em face da troca do ponto de interconexão à Internet.

No modelo base do IP móvel, cada estação da rede é tratada como uma estação isolada. Caso uma das estações não tenha acesso direto ao ponto de interconexão, ou ponto de acesso, essa estação perde a conectividade com a rede, como visto na Figura 2.17(a). Propostas como o NEMO (*NEtwork MObility*) [McCarthy et al. 2009] e as suas variantes NEMO+ [McCarthy et al. 2008b], Light-NEMO+ [Sabeur et al. 2006] e MANEMO (*MAnet* NEMO) [McCarthy et al. 2009, McCarthy et al. 2008a] vêm sendo realizadas para possibilitar que mesmo os nós sem acesso direto ao ponto de interconexão permaneçam conectados e com acesso à Internet. Assim, a conectividade pode ser garantida não somente para uma estação, mas para uma rede inteira. De maneira geral, o NEMO define que uma das estações da rede móvel deve ser escolhida para funcionar como um roteador que provê acesso à Internet a todas as outras estações da rede móvel. Assim, as estações que não possuem acesso direto ao ponto de interconexão devem encaminhar os seus pacotes até esse roteador, denominado roteador móvel (*Mobile Router* - MR). A Figura 2.17(b) ilustra simplificadamente o funcionamento das propostas para mobilidade de rede, na qual o nó móvel A da Figura 2.17(a) exerce a função de roteador móvel.



(a) Modelo básico do IP móvel. (b) Mobilidade de rede.

Figura 2.17. Conectividade em redes sem-fio móveis.

O IETF (*Internet Engineering Task Force*) criou o grupo de trabalho NEMO para trabalhar no tema de mobilidade de rede [Perera et al. 2004]. O grupo de trabalho NEMO é responsável pelo desenvolvimento e padronização dos protocolos para suporte da mobilidade de rede sobre o IP. O projeto inicial é fortemente baseado no IP móvel. Dessa forma, o NEMO pode ser implementado como uma extensão do IP móvel, na qual as funcionalidades definidas pelo IP móvel para uma estação são transferidas para o roteador móvel. Como consequência, as alterações após eventuais mudanças de ponto de interconexão com a Internet são tratadas apenas pelo roteador móvel. Tais alterações, como mudança de endereço IP (*Care-of-Address* - CoA) e restabelecimento do túnel com

o agente domiciliar, tornam-se transparentes aos outros nós da rede móvel. Com exceção do roteador móvel, todos os outros nós da rede mantêm suas configurações constantes. O NEMO também utiliza agentes domiciliares como o IP móvel para manutenção da conectividade. No NEMO, os roteadores móveis também enviam mensagens de atualização aos seus agentes domiciliares. Entretanto, os agentes domiciliares associam o endereço do roteador móvel na rede estrangeira (CoA) a um prefixo de rede ao invés de apenas a um endereço IP. Assim, todos os pacotes recebidos pelo agente domiciliar destinados a um dos nós da rede móvel são encapsulados e encaminhados através do túnel até o roteador móvel. O roteador móvel, por sua vez, desencapsula os pacotes recebidos e os encaminha até os seus respectivos destinos na rede móvel. Além disso, o roteador móvel também encaminha para a Internet todos os pacotes oriundos da rede móvel. Caso haja filtragem de egresso na rede estrangeira, os pacotes são encapsulados e encaminhados através do túnel no sentido reverso até o agente domiciliar. O agente domiciliar, por fim, desencapsula os pacotes e os encaminha via roteamento IP até os nós correspondentes na Internet. A Figura 2.18 ilustra o caminho seguido pelos pacotes usando o NEMO.

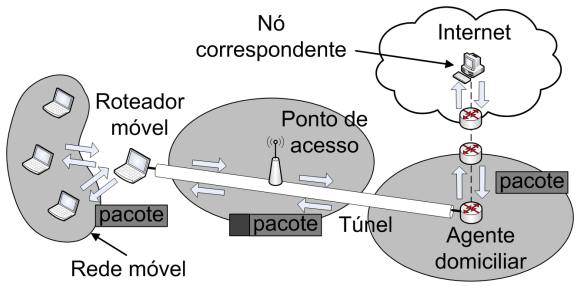


Figura 2.18. Encaminhamento de pacotes no NEMO.

O grupo de trabalho responsável pelo NEMO padronizou inicialmente um protocolo de suporte básico (NEMO *Basic Support Protocol* - NEMO BS) para o IPv6 e tem abordado diferentes problemas em outros documentos, inclusive a implementação do NEMO sobre o IPv4 [Devarapalli et al. 2005]. A sinalização utilizada pelo NEMO BS é uma extensão das mensagens definidas pelo IP móvel. As mensagens do NEMO BS possuem uma *flag* adicional que identifica se a mensagem foi enviada por um roteador móvel ou por um nó móvel. Essas mensagens são enviadas utilizando o cabeçalho de extensão de mobilidade, no caso do IPv6; ou mensagens de controle sobre o UDP, no caso do IPv4. As mensagens mais utilizadas pelo NEMO BS são as mensagens relativas à atualização de localização do roteador móvel (*binding updates*). As mensagens de atualização de associação com as redes estrangeiras e os respectivos reconhecimentos são utilizadas para notificar os agentes domiciliares do novo ponto de interconexão com a Internet (CoA). As mensagens de atualização contêm o novo CoA, a *flag* adicionada pelo NEMO BS e o prefixo da rede móvel. Este último, porém, é uma opção que depende do modo de operação do NEMO BS, que pode ser implícito ou explícito. No modo implícito, as mensagens de atualização não contêm o prefixo da rede móvel. Nesse caso, os agentes domiciliares descobrem o prefixo de alguma maneira não definida pelo protocolo de suporte básico. Já no modo explícito, o prefixo da rede é adicionado às mensagens de atualização. Uma vez recebida a mensagem de atualização, independente do modo, os agentes domiciliares enviam o reconhecimento positivo correspondente. A Figura 2.18

ilustra como o NEMO BS especifica o roteamento dos pacotes enviados e recebidos pelos nós da rede móvel. Os roteadores móveis possuem uma interface IPv6 ou IPv4 pertencente à rede móvel, interface de ingresso que pode ser configurada de maneira estática; e outra interface conectada à rede que oferece conexão de saída para a Internet, interface de egresso. A interface de egresso é registrada na rede domiciliar a partir da associação realizada com o endereço IP da rede estrangeira (CoA). Um túnel bidirecional é estabelecido entre o roteador móvel e o seu agente domiciliar. O agente domiciliar, ao invés de somente encaminhar o tráfego recebido destinado ao roteador móvel através do túnel, encaminha também todo tráfego do prefixo de rede associado ao roteador móvel.

O protocolo de suporte básico NEMO BS deve lidar com desafios relacionados com o ambiente de operação e com a forma de implementação escolhida para estender o IP móvel. Desafios relacionados com a segurança e com o desempenho devem ser considerados. Por exemplo, para garantir que a origem dos pacotes enviados através do túnel seja verdadeira, tanto o roteador móvel quanto o agente domiciliar devem checar se o endereço IP de origem dos pacotes é um endereço IP pertencente à faixa de endereços da rede domiciliar. Embora os desafios relacionados com a segurança sejam relevantes, muitos trabalhos na área visam aumentar o desempenho do NEMO. Um dos problemas frequentemente abordados é a ineficiência do roteamento. Como visto, caso um nó da rede móvel deseje se comunicar com um nó na Internet, chamado de nó correspondente, todos os seus pacotes devem ser enviados primeiramente ao seu agente domiciliar. O agente domiciliar, então, encaminha os pacotes até o nó da Internet, o que torna o roteamento sub-ótimo. O problema pode ser agravado considerando que um único roteador móvel pode ter associado a ele mais de uma rede móvel. Fundamentalmente, nada impede que coexistam mais de uma rede móvel e que o roteador móvel de uma rede utilize o roteador móvel de outra rede para se comunicar com a Internet. Nesse caso, o NEMO é conhecido como *Nested* NEMO [McCarthy et al. 2008b] e as comunicações podem ser bastante ineficientes visto que todo tráfego enviado pelos roteadores móveis devem sempre passar pelos seus agentes domiciliares. A Figura 2.19 ilustra o problema.

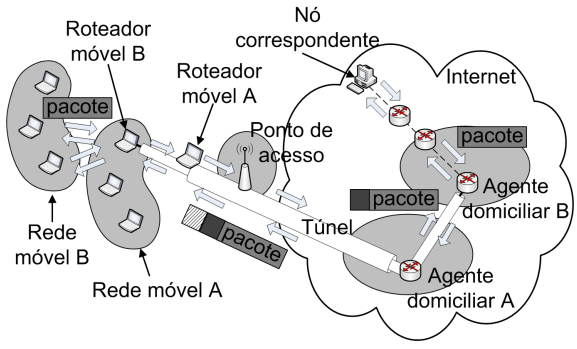
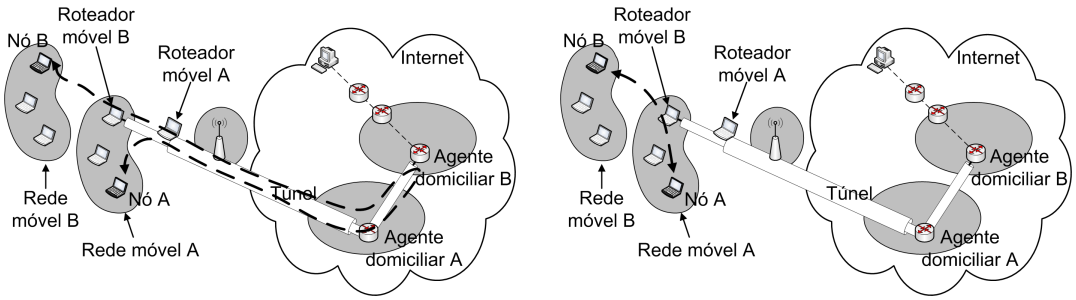


Figura 2.19. Problema conhecido como *Nested* NEMO.

Propostas para aumentar a eficiência do roteamento utilizam caminhos diretos entre os nós da rede móvel e os nós da Internet e até mesmo entre os nós em redes móveis diferentes. Algumas propostas analisam a possibilidade do uso de roteamento pela fonte para evitar que os pacotes passem sempre pelos agentes domiciliares. O NEMO+ [McCarthy et al. 2008b] propõem três protocolos para tornar as comunicações

envolvendo nós de redes móveis vizinhas mais eficientes. Esse tipo de comunicação ocorre quando uma rede móvel não possui acesso direto à Internet, e assim utiliza a rede vizinha para encaminhar o seu tráfego, e quando os nós móveis de diferentes redes desejam se comunicar. O primeiro protocolo proposto pelo NEMO+ é o TD (*Tree Discovery*). Esse protocolo é utilizado para auxiliar os roteadores móveis a escolherem qual dos roteadores móveis vizinhos oferece o melhor caminho até a Internet. A escolha é baseada nos anúncios IPv6 de descoberta de vizinhança enviados pelos vizinhos. Cada anúncio contém o caminho utilizado até o ponto de interconexão com a Internet obtido através de mensagens ICMPv6. Esse caminho representa um ramo da árvore formada desde o ponto de interconexão da rede (raiz) e os roteadores móveis vizinhos (folhas). Ao receber o anúncio, cada roteador escolhe o ramo da árvore que mais lhe convém se associar. Além disso, as mensagens de descoberta de vizinhança podem ainda ser utilizadas para evitar a formação de laços de roteamento (*loops*). Outro protocolo, o NINA (*Network In Node Advertisement*) é utilizado para anunciar aos roteadores móveis localizados em posições mais próximas ao ponto de interconexão os prefixos das sub-redes associados a cada roteador móvel. Esses anúncios possibilitam que as comunicações entre nós móveis associados a diferentes redes da mesma árvore possam se comunicar sem a necessidade do uso dos agentes domiciliares. A Figura 2.20 ilustra o caminho percorrido pelos pacotes do nó B na rede móvel B até o nó A na rede A com e sem o uso do protocolo NINA.



(a) Encaminhamento sem o protocolo NINA. (b) Encaminhamento com o protocolo NINA.

Figura 2.20. Encaminhamento com e sem o uso do protocolo NINA.

O último protocolo, chamado RRH (*Reverse Routing Header*), é usado para tornar o roteamento de pacotes enviados para a Internet também mais eficiente. No RRH, todo roteador móvel antes de encaminhar um pacote para a Internet atualiza o endereço IP de origem do pacote. O procedimento de atualização estabelece que o endereço IP de origem deve ser o do roteador móvel que encaminha o pacote e o endereço de origem anterior deve ser armazenado em uma lista no cabeçalho. Essa lista contém o endereço IP de todos os roteadores móveis já atravessados. Como o endereço IP de origem do pacote que é recebido pelo ponto de interconexão é do último roteador móvel, o pacote será somente encaminhado até o agente domiciliar desse último roteador antes de ser enviado até o nó correspondente. O RRH define que o cabeçalho contendo a lista de endereços IP de todos os roteadores móveis deve ser adicionado pelo agente domiciliar no caminho reverso para que o pacote possa ser entregue ao nó móvel de origem. Caso contrário, o pacote chegaria ao último roteador móvel e não seria possível identificar a verdadeira origem do pacote. O Light-NEMO+ [Sabeur et al. 2006] possui uma abordagem um pouco

diferente do protocolo RRH. No Light-NEMO+, o último roteador armazena um identificador do fluxo, semelhante a um *cookie*, e encaminha o pacote até o nó correspondente na Internet. Os pacotes cuja fonte ou destino estejam na rede móvel também carregam o *cookie* de identificação do fluxo. Assim, todo pacote vindo da Internet recebido pelo último roteador móvel possui uma identificação que pode usada para mapeamento do pacote para o fluxo correspondente. Note que o protocolo RRH usava roteamento pela fonte para identificar o destino ou a origem do pacote na rede móvel. Já o Light-NEMO+ utiliza um *cookie* e, portanto, dispensa o uso do roteamento pela fonte.

O MANEMO (*MAnet* NEMO) [McCarthy et al. 2008a] é um protocolo que tem por objetivo garantir que os nós móveis de uma rede ad hoc possam sempre ser alcançados de qualquer ponto na Internet. O roteador móvel que executa o protocolo MANEMO possui sua interface de ingresso configurada conforme a rede estrangeira visitada e a interface de egresso configurada conforme a rede ad hoc a qual faz parte. A interface de egresso executa um protocolo de roteamento ad hoc, p. ex. o OLSR [Clausen et al. 2001] configurado para se anunciar como *gateway* para a Internet. A arquitetura unificada MANEMO (*Unified MANEMO Architecture* - UMA) [McCarthy et al. 2009] é uma proposta para unificar os protocolos NEMO e assim garantir conectividade permanente aos nós da rede ad hoc. A UMA define a maneira pela qual os diferentes nós móveis conectam a Internet, via acesso direto a pontos de interconexão ou via outras redes móveis, e como os túneis entre os diferentes agentes domiciliares são estabelecidos. É responsabilidade dos agentes domiciliares identificarem os túneis criados para comunicação com cada um dos nós móveis assim como o restabelecimento da comunicação caso haja mudanças no posicionamento desses nós. A alteração do posicionamento dos nós móveis pode levar a mudanças do roteador móvel e, consequentemente, dos túneis previamente estabelecidos. A arquitetura UMA deve lidar com a dinamicidade da rede.

2.4.4. Múltiplos caminhos

A Internet atual é baseada em algoritmos de roteamento de caminho único. Dessa forma, os protocolos intradomínio e, principalmente, os interdomínio representados pelo BGP anunciam aos seus vizinhos apenas uma única opção de caminho para cada destino da rede. Uma maneira de contornar essa limitação e implementar os múltiplos caminhos é através de roteamento pela fonte. A definição de rotas a priori é uma solução para utilizar caminhos diferentes do padrão anunciado pelos protocolos de roteamento. Um exemplo de tecnologia que usa roteamento pela fonte é o PNNI (*Private Network-to-Network Interface* ou *Private Network Node Interface*) utilizado no ATM. O PNNI divide a topologia da rede em diferentes níveis hierárquicos e define por quais roteadores de cada nível os pacotes devem ser encaminhados [Kaur et al. 2003].

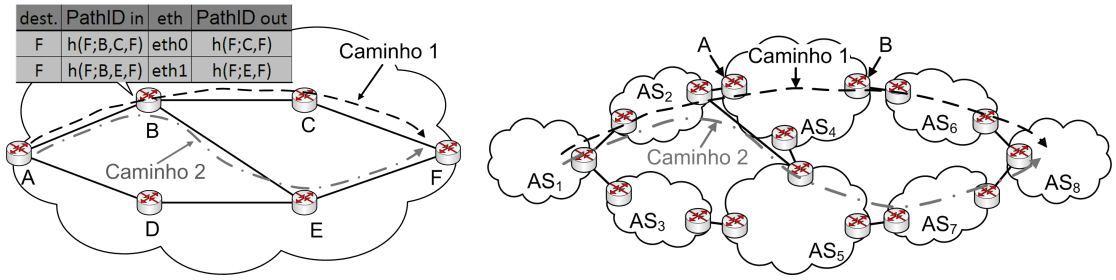
O arcabouço BANANAS [Kaur et al. 2003] é outra proposta que utiliza roteamento pela fonte. O BANANAS usa o conceito de `PathIDs` para identificação de caminhos. O `PathID` é a saída de uma função *hash* dos identificadores dos vértices e dos enlaces que compõem o caminho entre dois nós quaisquer em uma rede. Os identificadores dos vértices, p. ex. endereços IP, dos enlaces e de ASes são globalmente conhecidos e, portanto, o `PathID` também. O conceito de "globalmente conhecido" varia se o escopo for intradomínio (identificadores de enlaces e vértices) ou interdomínio (identificadores de ASes). No caso intradomínio, o conhecimento global é conquistado por algoritmos

de roteamento baseados em estado do enlace, já no caso interdomínio, pelos vetores de caminho do BGP. Para evitar colisões entre PathIDs, o PathID utilizado é estendido a uma tupla formada pelo endereço IP do destino e o resultado da *hash* calculado. Os PathIDs são adicionados a cada pacote para serem usados posteriormente durante o encaminhamento.

O BANANAS considera o roteamento intra e interdomínio. Vale mencionar que o BANANAS pode operar mesmo se apenas uma parte dos roteadores implementar o arcabouço proposto. Para isso, o cálculo do roteamento é realizado com restrições, pois considera como possíveis caminhos aqueles que passam por roteadores que implementam o BANANAS. No intradomínio, o roteador fonte envia os pacotes pelos múltiplos caminhos encontrados, como ilustrado na Figura 2.21(a). Essa figura mostra a tabela de encaminhamento do roteador B e o encaminhamento realizado para pacotes de mesma origem e destino, mas que seguem caminhos diferentes na rede. A função *hash* usada para o cálculo dos PathIDs é denotada por $h(.)$ na figura. Os pacotes carregam o PathID correspondente ao caminho seguido. Um roteador intermediário que implementa o BANANAS utiliza para o encaminhamento, ao invés da tupla prefixo do endereço de destino, próximo salto e interface de saída; a tupla prefixo do endereço de destino, PathID de entrada, interface de saída e PathID de saída. O PathID de entrada é a *hash* de todos os identificadores dos roteadores a partir do roteador atual até o destino e o PathID de saída é a saída da função *hash* dos identificadores desde o próximo salto até o destino. Ao receber um pacote, um roteador busca a entrada correspondente em sua tabela de roteamento baseado no prefixo do endereço de destino e no PathID de entrada. Antes de encaminhar o pacote, o roteador substitui o PathID do pacote pelo PathID de saída. Assim, o próximo roteador no caminho pode repetir o mesmo procedimento de encaminhamento. No BANANAS, todos os roteadores devem executar um algoritmo que calcule múltiplos caminhos já que o arcabouço foi proposto com essa finalidade. Mesmo os roteadores intermediários devem conhecer todos os possíveis caminhos a partir dele próprio até os possíveis destinos da rede para encaminhar pacotes seguindo a rota escolhida pelas fontes. O efeito do armazenamento dos múltiplos caminhos é o aumento das tabelas de roteamento. Entretanto, o BANANAS aborda esse compromisso utilizando técnicas de codificação para armazenamento de informações compactadas.

O roteamento interdomínio é uma abstração do roteamento intradomínio substituindo roteadores por ASes. O funcionamento é bastante semelhante se considerado o papel dos roteadores de borda de entrada e saída do AS semelhante ao papel das interfaces de entrada e saída de um roteador. Um PathID é adicionado aos pacotes, nesse caso chamado de e-PathID, no qual os identificadores são os identificadores dos sistemas autônomos do caminho. Um roteador de borda de entrada de um AS ao receber um pacote examina o prefixo do endereço de destino e encaminha até o roteador de borda de saída correspondente. Para isso, o endereço IP do roteador de borda de saída é utilizado como endereço de destino do pacote. O endereço de destino original é armazenado em uma estrutura em pilha na qual o endereço do roteador de borda de saída é inserido no topo. O roteador de borda de saída retira o seu endereço do topo da pilha e reinsere o endereço IP original do destino no pacote. O roteador de borda de entrada do próximo AS repete o procedimento de empilhamento de endereços IP. O PathID utilizado pelo roteamento em cada AS deve ser calculado baseado no caminho intradomínio escolhido. O uso de

múltiplos caminhos interdomínio deve lidar com questões relacionadas a políticas e acordos entre ASes. Essas questões não afetam o encaminhamento intradomínio que pode ser realizado sempre através de múltiplos caminhos quando disponíveis. A Figura 2.21(b) ilustra um exemplo de múltiplos caminhos interdomínio entre fontes no AS_1 e destinos no AS_8. Ainda na figura, os roteadores A e B representam o roteador de borda de entrada e saída, respectivamente, do AS_4 com relação ao fluxo do caminho 1. Logo, um pacote seguindo o caminho 1 utiliza como endereço IP de destino o endereço do roteador B após ser encaminhado por A.



(a) Roteamento intradomínio. (b) Roteamento interdomínio.

Figura 2.21. Encaminhamento de pacotes conforme o arcabouço BANANAS.

Redes sobrepostas também podem ser utilizadas para prover múltiplos caminhos. A vantagem dessa abordagem é não exigir modificações dos equipamentos do núcleo da rede. Os múltiplos caminhos são escolhidos na camada sobreposta e o encaminhamento dos pacotes segue o processo tradicional da Internet. A arquitetura RON (*Resilient Overlay Network*) [Andersen et al. 2001], inicialmente proposta para recuperação de falhas de rede, pode ser utilizada também para proporcionar múltiplos caminhos na Internet. Entretanto, uma das principais desvantagens do seu uso, assim como do roteamento pela fonte, é a redução do controle das rotas escolhidas por parte dos ASes de trânsito. É interessante para esses ASes escolherem os vizinhos para engenharia de tráfego e estabelecimento de acordos comerciais. O protocolo MIRO (*Multipath Interdomain ROuting*) [Xu e Rexford 2006] oferece maior flexibilidade aos ASes de trânsito e, por isso, não usa nem redes sobrepostas nem roteamento pela fonte. O MIRO baseia-se na negociação entre ASes vizinhos para uso de múltiplos caminhos. Embora possivelmente os múltiplos caminhos existam, cada AS anuncia apenas o caminho que mais lhe convém por questões de políticas, implementação dos protocolos e escalabilidade. Entretanto, o MIRO argumenta que um determinado AS deve requisitar os múltiplos caminhos, caso tenha interesse. Assim, problemas de escalabilidade são contidos e a implementação em toda Internet pode ser feita de maneira gradual. Um AS que não use o MIRO não responde às requisições por caminhos alternativos. O encaminhamento, nesse caso, acontece como na Internet atual. A Figura 2.22 ilustra o procedimento de negociação entre os roteadores A e B por caminhos alternativos. O roteador A pertencente ao AS_2 não deseja encaminhar o seu tráfego através do caminho padrão anunciado pelo AS_4 que é através do AS_6. O AS_4 oferece um caminho alternativo através do AS_5 que é aceito pelo AS_2. Os ASes podem fazer requisições para mais de um AS vizinho para obter mais caminhos alternativos e os ASes podem anunciar caminhos alternativos conhecidos caso atenda uma requisição recebida. Assim, um AS pode utilizar os caminhos alternativos dos seus vizinhos.

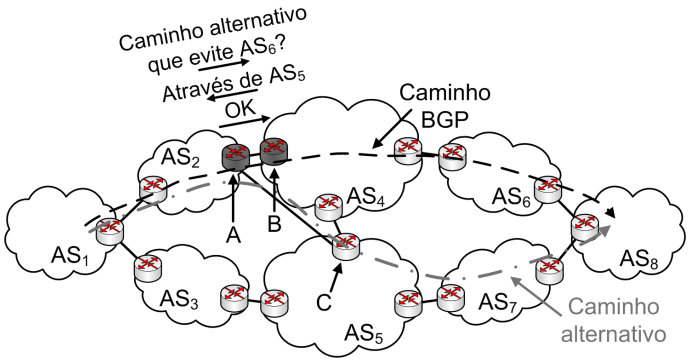


Figura 2.22. Negociação do MIRO para conhecimento de caminhos alternativos.

No MIRO, os caminhos descobertos pelo BGP padrão continuam a ser utilizados e os alternativos são usados através de procedimentos de tunelamento [Xu e Rexford 2006]. Logo, os pacotes que são enviados via caminho alternativo são enviados através de um túnel formado entre roteadores dos ASes vizinhos para garantir que o caminho alternativo seja, de fato, utilizado. O identificador do túnel é enviado pelo AS mais próximo ao destino para o AS mais próximo à fonte após o término do procedimento de negociação. Na Figura 2.22, um túnel entre os roteadores A e C é formado para encaminhamento de pacotes pelo caminho alternativo. Os caminhos alternativos podem ser utilizados para divisão de tráfego, no qual o tráfego mais prioritário utiliza o caminho alternativo, e para balanceamento de carga. A maneira como o tráfego é dividido depende das políticas adotadas pelos ASes e pelos acordos entre eles.

O processo de tunelamento pode ser chamado também de encaminhamento através de pontos de deflexão [Wetherall 2006, He e Rexford 2008]. O ponto de deflexão é o roteador para o qual o caminho é desviado. No exemplo da Figura 2.22 o ponto de deflexão é o roteador C visto que o caminho BGP convencional levaria o tráfego através do roteador B no AS4. Após o ponto de deflexão, o tráfego pode seguir o caminho BGP convencional caso nenhum caminho alternativo adicional seja usado. O emprego de pontos de deflexão pode ocorrer também nas redes sobrepostas. Em casos extremos, o ponto de deflexão poderia ser a estação de um usuário em outro AS que estivesse participando do roteamento. A participação de usuários pode impactar na escalabilidade da rede, mas evita que modificações sejam feitas na rede. O ponto de deflexão pode ser escolhido pela fonte do tráfego. Entretanto, independente da situação, o destino do túnel deve estar ciente do tunelamento para encaminhar o tráfego ao destino correto. Outra maneira de usar caminhos alternativos sem o uso de tunelamento é através da inserção de rótulos (*tags*) nos pacotes [Motiwala et al. 2008]. Um usuário final que queira que os seus pacotes sejam encaminhados através de caminhos alternativos deve inserir um rótulo no pacote referente ao caminho desejado. Os rótulos podem variar de acordo com as propriedades do caminho alternativo. Caso um roteador não reconheça o rótulo ou não implemente o sistema, ele pode encaminhar o pacote pelo caminho padrão.

2.4.5. Escalabilidade na Internet

Muitas propostas para interconexão de redes na Internet do futuro possuem impacto direto no número de entradas nas tabelas de roteamento. Por exemplo, o uso de

múltiplos caminhos requer mais de uma entrada nas tabelas para um mesmo destino. Já o uso dos múltiplos domicílios requer que uma rede seja identificada por mais de uma faixa de endereços possivelmente disjuntos prejudicando a agregação. O uso de identificadores planos, como os usados nas técnicas de *Loc/ID Split* e roteamento plano, também dificulta a agregação de rotas já que o espaço de endereçamento não é organizado hierarquicamente. Os identificadores planos, em especial, ainda aumentam o espaço de endereçamento da Internet o que pode tornar o problema da escalabilidade ainda mais grave. O aumento do espaço de endereçamento também pode ser uma consequência do uso do IPv6 se os problemas de agregação persistirem. Além do número de entradas nas tabelas de roteamento, outro problema que afeta a escalabilidade na Internet é o número de mensagens de atualização de roteamento enviadas principalmente por ASes de borda. Essas mensagens são enviadas para todos os outros ASes e o número pode aumentar devido a configurações mal feitas ou a ações maliciosas [Massey et al. 2007, Jen et al. 2008].

O aumento acelerado do número de entradas nas tabelas de roteamento pode impactar significativamente a capacidade de armazenamento da maioria dos roteadores atuais. Tal impacto pode causar inconsistências entre as tabelas, ou até mesmo, problemas de funcionamento dos equipamentos de rede [Kim et al. 2009]. Já o problema do número excessivo de mensagens de controle pode causar sobrecarga de tráfego e instabilidades nas tabelas de roteamento. Uma das propostas mais simples investigadas para conter ambos os problemas é a redução da flexibilidade de faixas de endereços disjuntas para múltiplos domicílios. A redução da flexibilidade utiliza duas estratégias. A primeira é a eliminação da possibilidade das redes de acesso utilizarem faixas de endereços diferentes da faixa de endereço dos seus ISPs. A segunda é a eliminação da possibilidade das redes de acesso utilizarem faixas de endereços desagregadas das faixas recebidas dos seus ISPs [Jen et al. 2008]. Nessa proposta, as redes de acesso somente podem utilizar faixas de endereços pertencentes às faixas dos seus provedores. Ainda, as estações das redes de acesso que são multidomiciliadas devem receber múltiplos endereços IPs, onde os endereços pertencem à faixa dos ISPs diretamente conectados. Essa limitação permite que cada ISP anuncie prefixos de rede agregados. A Figura 2.23 ilustra esse tipo de abordagem. Entretanto, uma desvantagem dessa abordagem é que os administradores das redes de acesso devem concordar em utilizar apenas faixas de endereços dos seus provedores diretos, já que hoje já existe a possibilidade de usar faixas de endereços independentes do provedor. Como consequência, a redução das tabelas de roteamento na DFZ torna-se dependente dos interesses dos administradores. Essa dependência não existe na proposta de separação de endereços.

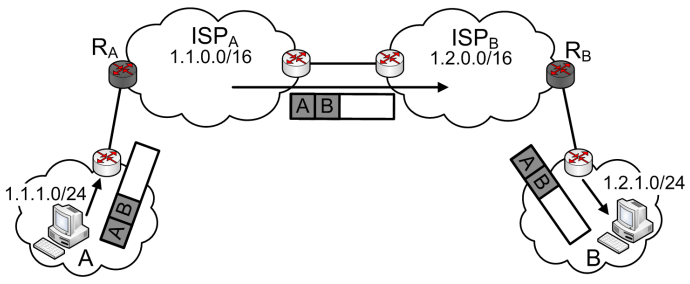


Figura 2.23. Encaminhamento em redes de acesso com sub-faixas de endereço de ISPs.

Uma proposta mais flexível para aumentar a escalabilidade na Internet é separar o espaço de endereçamento em Endereços Globalmente Roteáveis e Endereços Globalmente Entregáveis [Massey et al. 2007, Jen et al. 2008]. Os Endereços Globalmente Roteáveis são formados por endereços presentes nas tabelas de roteamento da DFZ e que são apenas alcançáveis na DFZ. Em oposição aos Endereços Globalmente Roteáveis, os Endereços Globalmente Entregáveis são os endereços de redes nas bordas da Internet que devem ser únicos e alcançáveis de qualquer lugar. Entretanto, os Endereços Globalmente Entregáveis não devem estar presentes nas tabelas da DFZ. A ausência de endereços de redes de borda nos roteadores da DFZ diminui o número de entradas e o número de prefixos anunciados pelo protocolo de roteamento interdomínio. Estimativas apontam que a eliminação dos prefixos de redes das bordas do roteamento interdomínio reduz o tamanho das tabelas e frequência de atualizações em até uma ordem de magnitude [Massey et al. 2007]. Além disso, um efeito indireto da separação de endereços é a possibilidade de emprego incremental de endereços diferentes do IPv4. Por exemplo, os endereços da borda podem ser IPv6 enquanto os do núcleo da rede podem ser IPv4.

A separação de endereços em dois tipos demonstra a divisão entre endereços de redes de ISPs, compostas tipicamente de roteadores de trânsito, e redes de acesso nas bordas. Assim, faixas de Endereços Globalmente Roteáveis são alocadas aos ISPs para que os diferentes ISPs sejam capazes de se comunicar. Um dado importante é que o número de ISPs tem sido estável em comparação ao número de redes de acesso [Massey et al. 2007]. Como consequência, as tabelas que contêm Endereços Globalmente Roteáveis não devem aumentar de tamanho rapidamente. Já os Endereços Globalmente Entregáveis são alocados para as redes de acesso. Esses endereços devem ser únicos para que cada rede de acesso ou estação seja identificada em toda Internet. Entretanto, como os provedores de serviço não conhecem os endereços das redes de acesso, este último não é globalmente roteável. Para que a correspondência seja possível, é necessário que haja um mapeamento dos dois tipos de endereços. Tal mapeamento é realizado nos roteadores de borda dos ISPs conectados às redes de acesso. Cada roteador de borda deve descobrir qual o mapeamento a realizar baseado no endereço de destino do pacote recebido. Após o mapeamento, o roteador encapsula o pacote e envia através do túnel formado até a roteador de borda da rede da estação do destino [Massey et al. 2007, Jen et al. 2008]. O roteador de borda da rede do destino desencapsula o pacote e entrega ao destino correspondente na rede de acesso.

O procedimento de mapeamento e encapsulamento [Jen et al. 2008] é ilustrado na Figura 2.24. Nessa figura, pacotes originados na estação A e destinados à estação B são enviados através de um túnel de R_A até R_B. A figura mostra também a divisão dos dois espaços de endereçamento existentes na proposta. É importante observar que os roteadores internos ao espaço de Endereços Globalmente Roteáveis não precisam conhecer os mecanismos de mapeamento e encapsulamento. Isso permite que a configuração deles permaneça a mesma que a atual, apenas com menos entradas nas tabelas de roteamento.

O uso do espaço de Endereços Globalmente Entregáveis permite que cada ISP agregue o maior número de faixas de endereços possível. Além disso, cada rede de acesso pode usar faixas de endereços independentes da faixa de endereços do ISP diretamente conectado, como visto na Figura 2.24. Essa característica facilita o uso dos múltiplos domicílios bem como a mudança das faixas de endereços utilizadas por cada rede de acesso. Entretanto, toda mudança realizada deve ser conhecida pelos roteadores responsáveis pelo

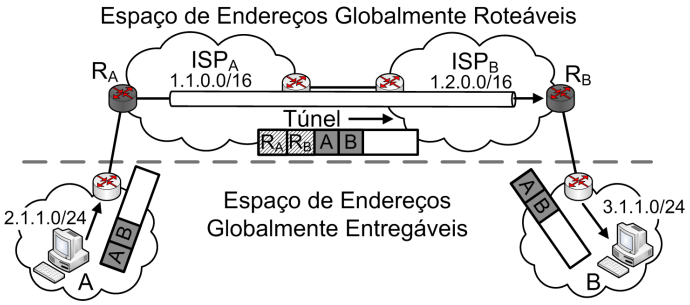


Figura 2.24. Encaminhamento utilizando dois espaços de endereçamento.

mapeamento para que esses continuem informando corretamente os outros roteadores na DFZ quais as redes de acesso que conhecem. As atualizações das informações de mapeamento podem acarretar problemas de escalabilidade uma vez que as atualizações são enviadas para em toda a DFZ.

Uma possibilidade para reduzir o número de mensagens de controle sobre o mapeamento é limitar as mensagens apenas para outros roteadores de borda utilizando técnicas, como por exemplo, o multicast. Outra possibilidade é o emprego de sistemas como o DNS para realizar o mapeamento. Todas as duas soluções reduzem a carga de controle, mas inserem outros problemas. Utilizar o multicast não é trivial na Internet e o uso de sistemas como o DNS pode inserir atrasos na resolução de endereços. Uma solução híbrida é proposta pela arquitetura APT (*A Practical Tunneling*) [Jen et al. 2009]. A arquitetura APT propõe que as informações de mapeamento sejam enviadas para todas as redes na DFZ. Entretanto, em cada uma das redes apenas um número reduzido de novos dispositivos de rede recebem tais informações. Esses dispositivos, denominados DM (*Default Mappers*), armazenam as tabelas completas com todos as informações sobre mapeamentos. Os roteadores de borda armazenam apenas as informações dos últimos mapeamentos realizados em uma memória cache. Logo, sempre que um pacote é recebido por um roteador de borda, ele verifica se o mapeamento específico está em sua memória cache. Caso esteja, o pacote é encapsulado e enviado, caso contrário, ele encaminha o pacote até o DM. O DM então trata o pacote como uma requisição para informação de mapeamento. Como resposta, o DM envia o mapeamento até o roteador de borda requisitante e, ao mesmo tempo, encapsula e encaminha o pacote em nome do roteador requisitante. Essa estratégia híbrida evita que os roteadores de borda armazenem tabelas completas, reduzindo o tamanho das tabelas utilizadas. Por outro lado, os DMs não fazem encaminhamento de muitos pacotes, reduzindo a necessidade de rapidez de encaminhamento. Essas características específicas permitem que os dispositivos sejam otimizados conforme a tarefa realizada.

O mapeamento deve lidar também com problemas de compatibilidade com equipamentos que ainda não separam o espaço de endereçamento [Vogt 2008]. Portanto, deve-se investigar maneiras de implementar a separação de endereços considerando a implementação gradual da proposta. Os roteadores de borda que não implementam a separação de endereços não realizam o mapeamento e o encaminhamento de pacotes. Além disso, eles podem não anunciar os prefixos das redes de acesso conhecidas. Logo, outros roteadores de borda descartam os pacotes recebidos cujo destino seja uma estação dessas redes de acesso que não tiveram seus prefixos anunciados. Isso ocorre porque os

roteadores de borda desconhecem o caminho até a rede do destino do pacote. Uma proposta para a adoção gradual da separação de endereços é o *Six/One Router* [Vogt 2008]. O *Six/One Router* é um protocolo para tradução de endereços entre estações em redes de acesso diferentes. A tradução é realizada no roteador de borda da rede de acesso de origem que mapeia o endereço de origem do pacote para um endereço de origem da DFZ. De maneira semelhante, o roteador de borda de origem realiza um procedimento de resolução de endereços para mapear o endereço de destino do pacote no endereço de destino na DFZ. Esse mapeamento é biunívoco, ou seja, cada endereço de uma estação em uma rede de acesso só pode ser mapeado em um endereço da DFZ. O endereço da DFZ pertence à faixa de endereços do ISP correspondente. O roteador executando o *Six/One Router* insere também uma extensão no cabeçalho do pacote para identificar os endereços de origem e destino das redes de acesso. Assim, quando o roteador de borda na rede de acesso do destino recebe o pacote, ele traduz os endereços para os endereços originais das redes de acesso e envia para a estação correspondente. O mapeamento é armazenado em memória cache para que o procedimento seja realizado mais rapidamente em pacotes seguintes. No caso de estações em redes de acesso legadas, ou seja, que não realizam a separação de endereços, o endereço da estação já é o próprio endereço da DFZ já que esses endereços têm que ser conhecidos globalmente. Portanto, caso um roteador de borda de origem realize a tradução, o endereço da DFZ da estação legada será o seu próprio endereço. No sentido reverso, caso uma estação legada seja a origem do pacote, o endereço de destino usado será o endereço da DFZ. Como o mapeamento é biunívoco, há somente uma estação associada a esse endereço da DFZ.

A Figura 2.25 ilustra o funcionamento do *Six/One Router* em caso de redes legadas. A estação A possui o seu endereço traduzido em um endereço do seu provedor de serviço. Entretanto, o endereço da estação B de destino é mantido, visto que a sua rede de acesso não realiza a separação de endereços. Logo, o endereço da estação B é conhecido globalmente e, por isso, pode ser mantido no pacote. A figura ilustra também a extensão incluída no cabeçalho do pacote pelo roteador R_A. Note que enquanto a rede de acesso da estação A não pode utilizar uma faixa de endereços do seu provedor (ISP_A), a estação B pode usar uma faixa do ISP_B.

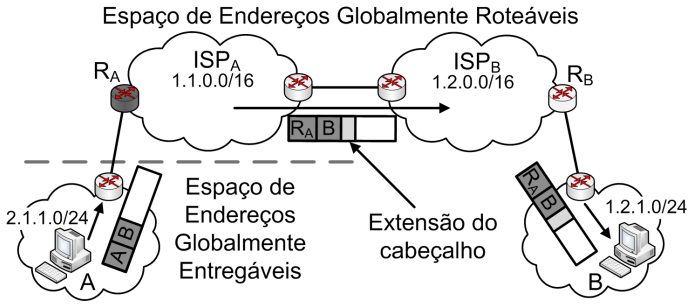


Figura 2.25. Funcionamento do *Six/One Router* em caso de redes de acesso legadas.

O HAIR (*Hierarchical Architecture for Internet Routing*) [Feldmann et al. 2009] é outra proposta para reduzir o número de entradas nas tabelas de roteamento da DFZ. Diferente das propostas de separação de endereços em dois níveis, o HAIR pode dividir os diferentes ASes em mais de dois níveis hierárquicos, onde o mais alto é composto

pelos principais ASes de trânsito da Internet e o mais baixo é composto pelas redes de acesso. Além disso, o HAIR utiliza técnicas para *Loc/ID Split* para possibilitar a mobilidade das estações. O encaminhamento de pacotes é realizado, portanto, baseado nos identificadores das estações de origem e destino que são mapeadas em endereços pela origem antes do envio dos pacotes. No HAIR, o endereço de origem codifica os identificadores de todos os roteadores de borda pelos quais o pacote deve passar desde sua origem até o nível hierárquico superior. Semelhantemente, o endereço de destino codifica os identificadores dos roteadores de borda desde o nível superior até o destino. O procedimento de encaminhamento é parecido com a opção *loose source routing* do IP. A diferença é que o caminho não é inserido no campo de opção do IP e é codificado nos endereços de origem e destino utilizados. O mapeamento do identificador do destino no endereço é realizado sob requisição. Para tal, assume-se que a estação de origem já conhece o identificador de destino e utiliza um sistema de resolução de identificação de estação em endereço. A Figura 2.26 mostra os endereços de origem e destino usados no pacote. Note que o endereço de origem é uma codificação do identificador da estação (estação A), do roteador do primeiro nível hierárquico acima (R_A) e do roteador do núcleo (R_{NA}). A divisão hierárquica dos ASes e o uso do roteamento pela fonte permitem que cada roteador conheça apenas as rotas até os roteadores de borda dentro do seu próprio nível hierárquico. Essa característica leva à redução do número de entradas nas tabelas de roteamento na DFZ.

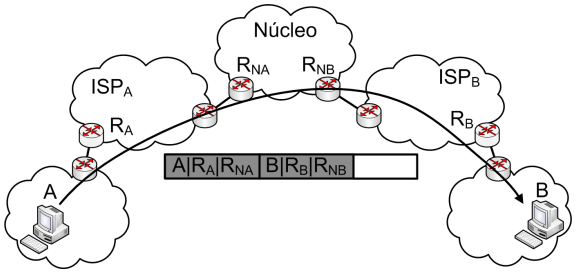


Figura 2.26. Funcionamento do HAIR.

Todos os trabalhos apresentados anteriormente reduzem o problema de escalabilidade na Internet através de novos esquemas de endereçamento. Entretanto, todas essas propostas necessitam alterar de certa forma a arquitetura atual de endereçamento, o que pode representar uma barreira para as suas implementações. Tendo em vista esse problema em curto e médio prazo, alguns trabalhos propõem a redução das tabelas de roteamento através de estratégias mais inteligentes de organização de prefixos ou de armazenamento em memória.

O ViAggre (*Virtual Aggregation*) [Ballani et al. 2008, Ballani et al. 2009] propõe distribuir a manutenção da tabela de roteamento intradomínio entre os roteadores de um ISP. Assim, cada roteador deve manter apenas parte da tabela completa de roteamento. O ViAggre divide o espaço de endereçamento em um conjunto de prefixos virtuais tal que cada prefixo virtual é maior que os prefixos agregados reais utilizados pelos roteadores. Por exemplo, todo o espaço de endereçamento conhecido por um ISP pode ser dividido em 128 prefixos virtuais /7 (0.0.0.0/7 até 254.0.0.0/7), no qual cada prefixo virtual corresponde a um conjunto de prefixos reais. Os prefixos virtuais não necessariamente possuem

correspondência com a topologia da rede, mas devem cobrir todos os possíveis prefixos reais para que não haja nenhuma rede inalcançável. As redes virtuais geradas a partir dos prefixos virtuais formam uma topologia que possui prefixos agregáveis e, portanto, tornam-se escaláveis. Para criar uma rede virtual, cada ISP define alguns roteadores para fazerem parte dessa rede. Esses roteadores mantêm rotas para todos os prefixos contidos no prefixo virtual. Tais roteadores são denominados pontos de agregação (*aggregation points*) para o prefixo virtual. Um ponto de agregação pode agregar mais de um prefixo virtual. Esse ponto de agregação deve apenas manter rotas para os prefixos contidos nos prefixos virtuais que ele mantém.

O encaminhamento de pacotes utilizando o ViAggre ocorre da seguinte maneira. Assim que o pacote entra na rede do ISP, esse pacote é encaminhado diretamente até o ponto de agregação mais próximo que mantém o prefixo virtual que engloba o prefixo do destino. Esse ponto de agregação possui uma rota para o prefixo do destino e, portanto, encaminha o pacote até o próximo ISP. O encaminhamento é realizado através de um túnel já que algum roteador no caminho pode desconhecer a rota escolhida e enviar o pacote de volta para o ponto de agregação. A divisão dos prefixos da Internet em prefixos virtuais é uma tentativa de rearrumar os prefixos para que eles se tornem agregáveis. Uma vez que os prefixos possam ser agregados, o número de entradas nas tabelas de roteamento diminui especialmente na DFZ.

Outras propostas atuam mais especificamente na organização da memória com relação aos tipos de entradas que devem ser armazenadas para reduzir o tamanho da tabela sem perder informação. Kim *et al.* [Kim et al. 2009] propõem armazenar em memória cache a tabela de roteamento apenas com as rotas usadas com maior frequência. As outras rotas são armazenadas em uma memória mais lenta e consultadas apenas em caso de ausência na cache. Caso um pacote seja recebido pelo roteador e uma rota não esteja disponível na memória cache, o roteador encaminha imediatamente o pacote em uma rota *default* e atualiza a sua memória cache baseado nas informações que possui em sua memória mais lenta. A atualização da cache segue a política de atualização adotada. O roteador que recebe o pacote enviado pela rota *default* conhece caminhos para todos os possíveis destinos. Esses roteadores podem ser projetados de maneira a armazenarem todas as possíveis rotas da rede. Além de armazenar parte da tabela de roteamento em cache, Kim *et al.* também propõem o uso de prefixos de mesmo comprimento. Através da análise de registros reais, os autores concluíram que prefixos /24 são os mais específicos possíveis que ainda não são filtrados pelos provedores por questões de segurança. Prefixos de comprimento menores, por exemplo /16, podem ser subdivididos em entradas com prefixos mais específicos que utilizam interfaces de saída diferentes da entrada com prefixo maior. Por exemplo, o prefixo 10.1.0.0/16 pode ter como interface de saída a eth0 enquanto o prefixo 10.1.1.0/24, a eth1. Caso um pacote destinado ao endereço 10.1.2.1 seja recebido, a entrada com prefixo maior é colocada em cache e o pacote é encaminhado pela eth0. Se em seguida um pacote com endereço de destino 10.1.1.1 for recebido, o pacote também é encaminhado pela eth0 ao invés da eth1 como definido pela entrada mais específica. A divisão em prefixos /24 é, portanto, uma tentativa de se evitar problemas durante o encaminhamento de pacotes ao utilizar parte das rotas em cache.

Organizar de maneira mais inteligente os prefixos nas tabelas de roteamento pode enfrentar dois desafios. O primeiro é que a solução proposta pode não ser definitiva.

Mudanças na arquitetura da Internet podem ser mais efetivas, porém mais difíceis de implementar sem incentivos consideráveis. Já o segundo problema é um compromisso verificado por Krioukov *et al.* [Krioukov et al. 2007] entre a redução das tabelas de roteamento e o aumento do tamanho médio das rotas. Krioukov *et al.* demonstram que, na Internet, o crescimento das tabelas de roteamento é logarítmico se forem utilizados algoritmos para compactação das tabelas. Entretanto, tal crescimento só é possível caso a Internet seja estática e os endereços estiverem relacionados com a topologia da rede. Se forem utilizados identificadores planos, o crescimento das tabelas de roteamento é polinomial, onde no melhor caso o crescimento é linear se os algoritmos de roteamento buscarem os caminhos mais curtos. Caso os caminhos encontrados possuam certa tolerância e não necessariamente sejam sempre os mais curtos, essa taxa de crescimento pode ser menor, no melhor caso.

2.4.6. Caminhos programáveis

A provisão da qualidade de serviço na Internet deve lidar com os diferentes requisitos das aplicações. Para possibilitar serviços diferentes do tradicional "melhor esforço", funcionalidades para visualização da topologia da rede bem como métodos sofisticados para projetos de novas aplicações são necessários. Essas novas técnicas podem ser utilizadas para mover funções das aplicações para a rede, já que o desempenho da rede pode variar [Clark et al. 2004]. Maneiras para monitorar o desempenho da rede, como movimentar as funções das aplicações e auxiliar o processo de escolha dos caminhos, podem ser feitas através da introdução de inteligência na rede. Tal inteligência é viabilizada através da aquisição de experiência e visão global. A inteligência é adquirida por agentes que atuam na rede e utilizam o conhecimento disponível, frequentemente obtido no Plano de Conhecimento [Clark et al. 2003], para tomar decisões sobre quais são os melhores caminhos conforme a aplicação. Além do emprego de agentes, a escolha do caminho pode também ser feita a partir do próprio usuário. Para tanto, arquiteturas que ofereçam ao usuário a oportunidade de escolher qual o caminho seguido pelo seu tráfego no nível de ASes devem ser desenvolvidas. As propostas nessa área podem ser classificadas de duas maneiras: orientadas a agentes ou orientadas a usuários.

As propostas orientadas a agentes estão na direção oposta à premissa fim-a-fim da Internet. A característica de inteligência nas bordas, por um lado, aumenta a simplicidade e flexibilidade da rede [Moreira et al. 2009]. Entretanto, por outro lado, dificulta o diagnóstico de falhas e implica em configurações manuais. Essas últimas consequências reduzem o desempenho da rede já que podem gerar configurações erradas e levar ao aumento do tempo de recuperação da rede. A inserção de agentes inteligentes na rede tem por objetivo reduzir os problemas acarretados por configurações manuais. Entretanto, o emprego dos agentes não deve prejudicar um dos principais pilares da Internet que é a simplicidade do núcleo [Clark et al. 2003].

O papel dos agentes na Internet é aumentar a autonomia da rede ao ponto de torná-la o mais independente possível da intervenção humana. Atualmente, o crescimento do número de nós, de usuários e da demanda por conectividade e banda passante têm tornado o gerenciamento das redes mais e mais complexo. O gerenciamento humano, por conseguinte, pode se tornar um limitante para o crescimento da Internet. O conceito de redes autonômicas defende o uso de técnicas que permitam à rede conhecer o contexto que está

inserida e o que lhe é solicitada. Assim, torna-se possível realizar ações sem a intervenção humana. As redes autonômicas também devem ter uma visão em alto nível dos objetivos da rede e das suas limitações para tomar decisões de configuração. Por fim, os agentes devem relatar o seu desempenho em alto nível para usuários ou administradores.

O conhecimento do tráfego encaminhado e dos requisitos dos usuários permite que os agentes tomem decisões no nível do roteamento. Tais decisões não necessariamente são regidas por resultados de algoritmos determinísticos. Isso ocorre porque o ambiente no qual as decisões são tomadas é altamente dinâmico e sujeito a conflito de interesses e a informações incompletas. Nesse caso, é importante manter uma base de dados que construa, reconcilie e mantenha os muitos aspectos de uma visão em alto nível do comportamento da Internet. Essa base de dados é separada em um novo plano denominado Plano de Conhecimento (*Knowledge Plane*) por Clark *et al.* [Clark et al. 2003]. O Plano de Conhecimento deve ser capaz de prover serviços aos outros elementos da rede como, por exemplo, qual a melhor configuração em um dado momento de operação da rede. No caso do roteamento, o Plano de Conhecimento é capaz de alterar ou estabelecer caminhos conforme requisitos de aplicações. Um dos desafios do Plano de Conhecimento é mensurar o quanto de informações ele deve possuir e qual o alcance delas. Especialmente no roteamento, projetar um Plano de Conhecimento que possua informações globais da Internet não é viável dada a massa de informações necessárias. Logo, o Plano de Conhecimento deve conhecer o tipo de informação que é mais útil em uma dada circunstância e deve usar técnicas distribuídas escaláveis para filtrar as observações conforme os interesses dos usuários.

Nas propostas orientadas a usuários, estes possuem maior liberdade para escolher o tipo de caminho pelo qual os seus pacotes são encaminhados. A arquitetura NIRA (*New Internet Routing Architecture*) [Yang et al. 2007, Yang 2003] é a principal representante desse tipo de propostas. A arquitetura NIRA tem como um dos seus objetivos estimular a competição entre os ISPs. Esse é o motivo pelo qual a arquitetura NIRA propõe que os usuários tenham oportunidade de escolher os caminhos no nível de ASes e não no nível de roteadores internos a um AS. A competição entre ISPs pode levar à redução dos custos do acesso à Internet para os usuários bem como estimular a introdução de valor agregado aos serviços. Dentre possíveis valores agregados poderia estar, por exemplo, caminhos com qualidade de serviço fim-a-fim.

A maneira mais trivial para os usuários definirem o caminho percorrido pelos pacotes é o uso de roteamento pela fonte. Entretanto, o roteamento pela fonte aumenta o tamanho dos cabeçalhos. A arquitetura NIRA utiliza um esquema de roteamento pela fonte que não introduz o caminho explicitamente no cabeçalho do pacote. Na NIRA, os endereços das estações são a concatenação do identificador da estação com os identificadores de todos os ASes até o núcleo da Internet. A parte do endereço que identifica os ASes deve ser uma concatenação dos prefixos de todos os domínios no caminho até os ASes de núcleo. A Figura 2.27 ilustra um exemplo, supondo que haja apenas um domínio entre a rede de acesso e o núcleo, o endereço da estação A é 1.1.1.1, que é uma concatenação do identificador da estação (1), do identificador do domínio intermediário (1.1.1.0/24) e do núcleo (1.1.0.0/16). Assim, o caminho do pacote fica definido até o núcleo da Internet a partir do endereço da estação. Em uma comunicação fim-a-fim, a maneira como os endereços de origem e de destino dos pacotes são organizados é suficiente

para que o caminho seja totalmente definido. Apenas os ASes de núcleo precisam executar um protocolo de roteamento para encaminhar os pacotes recebidos. Na Figura 2.27, o protocolo de roteamento define o melhor caminho a ser seguido pelos pacotes entre o roteador R_A e R_B, por exemplo. Como o número de ASes no núcleo é pequeno em relação ao número de ASes da Internet, esse protocolo não enfrenta problemas de escalabilidade.

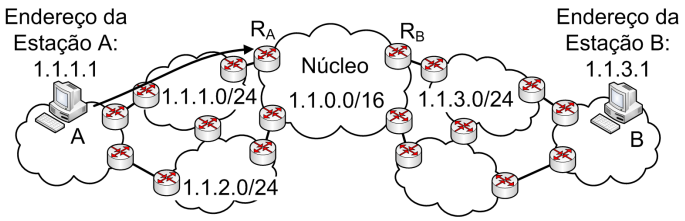


Figura 2.27. Estrutura de endereços utilizados na NIRA.

O encaminhamento de pacotes baseado tanto no endereço de origem quanto no destino dos pacotes é diferente do encaminhamento de pacotes da Internet que é baseado apenas no endereço de destino. Entretanto, a definição do caminho nos endereços dos pacotes ainda não é suficiente para prover liberdade de escolha aos usuários. Assim, é necessário um procedimento de descobrimento de rotas para que os usuários possam escolher, dentre os caminhos possíveis, o melhor dependendo dos seus interesses. A NIRA utiliza mecanismos de descoberta de caminhos antes do envio de pacotes. Para tal, a NIRA define dois protocolos para auxiliar a descoberta de caminhos: o TIPP (*Topology Information Propagation Protocol*) e o NRRS (*Name-to-Route Resolution Service*). O primeiro protocolo propaga aos usuários suas informações de endereços intradomínio levando em consideração possíveis acordos entre ASes. As informações correspondem ao caminho desde a estação até os ASes de núcleo. Na Figura 2.27, a estação A poderia escolher entre o endereço 1.1.1.1 e o endereço 1.1.2.1. Já o protocolo NRRS define um serviço de resolução de nomes para que cada estação conheça o endereço das outras estações da rede. Assim, a estação A descobre o endereço da estação B, 1.1.3.1, como visto na Figura 2.27. Os endereços de origem e destino são armazenados em memória cache para evitar buscas consecutivas pelos mesmos endereços, exceto em caso de falhas.

Outra proposta orientada a usuários é o *Pathlet Routing* [Godfrey et al. 2009]. No *Pathlet Routing*, cada ISP anuncia fragmentos de caminhos, denominados *Pathlets*, que podem ser concatenados pelo usuário para formar um caminho fim-a-fim. Um *Pathlet* é definido como uma sequência de nós considerados virtuais. Tais nós podem estar associados a todas as rotas conhecidas pelos roteadores da rede que pertencem, ou no caso mais simples, a apenas as rotas conhecidas por um único roteador. O caminho de um pacote é definido na fonte e inserido no cabeçalho do pacote. Logo, o *Pathlet Routing* utiliza roteamento pela fonte. Cada *Pathlet* é identificado por um identificador de encaminhamento (*Forwarding IDentifier* - FID) cujo significado é local ao nó virtual de origem do *Pathlet*. Uma vez que o roteamento é pela fonte, é importante apenas para o nó de origem do *Pathlet* conhecer para que nó encaminhar os pacotes recebidos. Ao encaminhar um pacote, o roteador altera a sequência de identificadores. A Figura 2.28 ilustra essas alterações. Na figura, cada seta pontilhada representa um *Pathlet* e os retângulos abaixo

de cada roteador representam os identificadores dos *Pathlets* contidos no cabeçalho do pacote naquele roteador. O pacote é originado no roteador A e destinado ao roteador E. Note que a cada salto os *Pathlets* atravessados são removidos do cabeçalho (ex. entre os roteadores A e B) e novos *Pathlets* são inseridos no cabeçalho do pacote caso um *Pathlet* atravessado seja composto por múltiplos saltos (ex. entre os roteadores B e C já que o *Pathlet* 2 é composto pelos *Pathlets* 7 e 1). *Pathlets* compostos são formados assim que um determinado roteador aprende múltiplos *Pathlets* consecutivos.

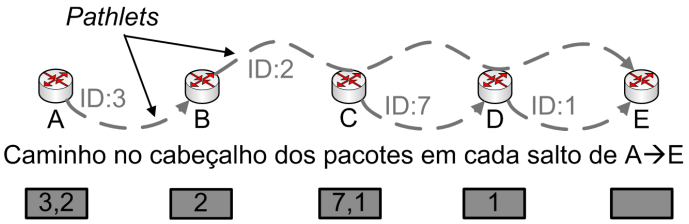


Figura 2.28. *Pathlet Routing*.

No *Pathlet Routing*, um determinado nó da rede dissemina os *Pathlets* conhecidos que traduzem as suas políticas. A disseminação é realizada através de vetores de caminhos no BGP. Os *Pathlets* recebidos pelos usuários são utilizados na escolha dos caminhos. Uma das vantagens do uso dos *Pathlets* é a economia no tamanho dos cabeçalhos. Na Figura 2.28, o roteamento pela fonte precisaria de cinco endereços IP enquanto com o uso do *Pathlet* precisa de no máximo dois identificadores (FIDs).

Uma questão importante é como incentivar os ISPs a disponibilizarem aos usuários a liberdade de escolha de caminhos. O modelo atual da Internet é baseado em acordos comerciais entre os provedores. Portanto, a manutenção desse modelo representa uma restrição à possibilidade de escolha dos usuários. A mudança do modelo, entretanto, deve ser atraente também para os ISPs. Alguns modelos de negócios para os provedores incluem acordos entre os usuários e os seus provedores de acesso e entre os usuários e os provedores de todo o caminho até o núcleo. Na primeira opção, os provedores conectados diretamente aos usuários fazem acordos com os provedores vizinhos para oferecerem diferentes possibilidades de caminhos para a escolha dos usuários. Os provedores tarifam os usuários através do monitoramento dos caminhos que os usuários escolheram. Já na segunda opção, os usuários têm a possibilidade de escolher todos os ASes no caminho. Embora essa última opção ofereça mais liberdade, ela pode sofrer problemas de adoção, visto que qualquer provedor de serviço segue algum tipo de política com os seus vizinhos.

Uma dificuldade que ambos os tipos de proposta, orientada a agentes e orientada a usuários, enfrenta é a aquisição de informações da rede. Tais informações são importantes para que tanto os agentes quanto os usuários possam embasar suas decisões. Por exemplo, aplicações como VoIP possuem restrições de atraso fim-a-fim. Essa métrica não é simples de ser obtida visto que necessita de sincronismo entre relógios. Além da obtenção de métricas, outro desafio é consolidar as informações e representar o estado atual da rede a partir dessas informações. O estado da rede deve retratar de maneira mais fiel possível dependendo das necessidades das aplicações. A arquitetura CONMan (*Complexity Oblivious Network Management*) [Ballani e Francis 2007] propõe uma interface para simplificar a obtenção de informações de protocolos. Os autores argumentam que uma

das grandes dificuldades para o gerenciamento de redes é que os protocolos e dispositivos exibem muitos detalhes de sua implementação, o que dificulta a obtenção dos dados. A arquitetura CONMan, portanto, propõe uma interface que inclui o mínimo necessário de informações específicas de protocolos e outros dispositivos para simplificar a obtenção de dados, e assim, melhorar o desempenho da rede a partir de um gerenciamento mais efetivo.

2.4.7. OpenFlow e a solução comutada

Atualmente, o Ethernet é uma tecnologia consolidada para redes locais e metropolitanas. O Ethernet é uma solução comutada e, portanto, não pode ser aplicada diretamente na Internet que é roteada. A simplicidade do Ethernet e a facilidade de configuração de redes, entretanto, vêm estimulando o emprego dessa tecnologia em redes de maior escala. Para isso, procedimentos como o de inundação de controle, utilizado para localização de estações, configuração automática de endereços através de DHCP (*Dynamic Host Configuration Protocol*) e resolução ARP (*Address Resolution Protocol*); e o emprego de árvores de espalhamento (*spanning tree*) para comunicação entre os nós devem ser revistos. Além disso, em uma rede comutada, cada comutador armazena a localização de cada destino na rede. Para contornar essas características e tornar a solução comutada mais escalável, trabalhos na área propõem soluções híbridas como o emprego de roteadores executando o IP para interconectar redes Ethernet e o emprego de VLANs (*Virtual LANs*) [Kim et al. 2008]. Outras propostas, como o Seattle [Kim et al. 2008], reduzem o número de estados por comutador e ainda evitam inundações. Para isso, o Seattle utiliza um sistema de localização de estações através de DHTs, e assim, dispensa o armazenamento de informações globais em cada um dos nós da rede.

Um das soluções comutadas que vem se destacando é o OpenFlow. O OpenFlow [McKeown et al. 2008, Mateo 2009] possibilita a programabilidade de elementos de rede, sejam eles comutadores ou roteadores. Essa característica permite programar redes e, como consequência, realizar experimentos isolados simultâneos para testes de novas propostas de roteamento e até mesmo alternativas ao IP. O OpenFlow permite que fluxos sejam definidos, bem como os caminhos seguidos por cada fluxo, sem interferir nos outros fluxos da rede. Os caminhos podem ser obtidos a partir de métodos oferecidos pelo OpenFlow. Tais métodos definem políticas para busca automática de caminhos que atendam requisitos desejados como largura de banda disponível e atraso restrito.

Uma rede convencional é composta por enlaces, elementos de processamento de pacotes (comutadores/roteadores) e estações finais. Nessas redes, cada elemento de processamento agrega funções de encaminhamento de pacotes e decisões de controle. No OpenFlow, entretanto, essas duas funções são separadas. A função de encaminhamento de pacote continua nos comutadores/roteadores. Porém, as decisões de controle são atribuídas a um novo elemento de rede, chamado de controlador. Nesse sentido, o OpenFlow altera a arquitetura convencional de redes ao introduzir o conceito de controladores. Os controladores separados foram propostos para que a programabilidade dos comutadores possa ser realizada externamente ao equipamento para que não haja dependência do fabricante. Essa característica facilita a adoção da tecnologia. Para que isso seja possível, o OpenFlow define uma interface e um protocolo para comunicação entre controladores e comutadores. Esse protocolo define mensagens para modificação das tabelas de enca-

minhamento, recepção e envio de pacotes, busca de estatísticas da rede e informações do equipamento. Essas mensagens são utilizadas para realizar determinadas funções, dentre as quais a mais importante para prover programabilidade é a modificação das tabelas de encaminhamento. A arquitetura de uma rede OpenFow é ilustrada na Figura 2.29.

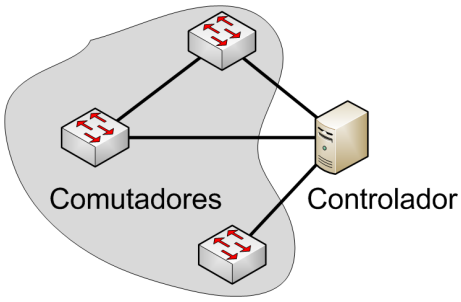


Figura 2.29. Arquitetura de uma rede OpenFlow.

A programabilidade do OpenFlow é obtida através da possibilidade de alterar a tabela de encaminhamento, que no caso da comutação, é chamada de tabela de fluxos. A tabela de fluxos define a ação associada a cada fluxo e estabelece o tipo de processamento a ser realizado sobre aquele fluxo. Cada fluxo é identificado na tabela a partir de informações contidas no seu cabeçalho. Informações como endereços MAC, endereços IP, portas TCP etc. são utilizadas para identificação. Uma vez identificado o fluxo, o OpenFlow executa a ação relacionada a ele como descrito na própria tabela. A presença de informação além de informações de camada de enlace pode ser usada por um comutador para que ele se comporte como um roteador ou como um equipamento de camada acima da camada de transporte. Da mesma maneira, um fluxo pode ser apenas identificado pelo cabeçalho da camada de enlace. Isso facilita o teste de soluções que não utilizam o IP como protocolo de rede. O OpenFlow não requer que o cabeçalho dos pacotes sigam um determinado formato. Ele requer apenas que o formato utilizado possa ser identificado na tabela de fluxos. Essa característica facilita a experimentação de novos protocolos para a Internet do Futuro.

Os pacotes recebidos que não possuem entradas correspondentes na tabela de fluxos são enviados ao controlador que decide a ação a ser realizada nesse pacote. Portanto, sempre que um pacote recebido não encontrar uma porta comutada de saída já estabelecida, ele é encaminhado para o controlador através do canal de comunicação. Esse canal utiliza SSL (*Secure Sockets Layer*) para garantir autenticação e confidencialidade, já que o protocolo definido pelo OpenFlow faz mudanças na tabela de fluxos dos comutadores. Após a análise do controlador, uma entrada é adicionada ou não na tabela de fluxos do comutador correspondente. Caso uma entrada seja adicionada, os próximos pacotes são encaminhados sem a necessidade de passar primeiro pelo controlador.

Uma rede OpenFlow é composta de um ou mais comutadores OpenFlow, um ou mais controladores e um canal seguro para comunicação. Cada comutador OpenFlow possui sua tabela de fluxos enquanto o canal seguro é usado para comunicação entre os comutadores e os controladores. Normalmente, em uma rede OpenFlow há comutadores e apenas um controlador. Como consequência, o OpenFlow possui controle centralizado. Essa centralização representa um compromisso entre escalabilidade e facilidade

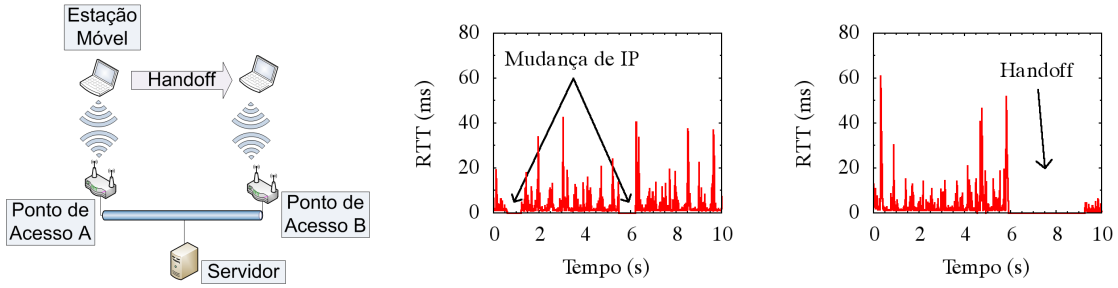
de gerenciamento. Por um lado, a centralização requer que todas as decisões de controle e gerenciamento da rede sejam executadas por apenas um elemento. Esse elemento pode representar um ponto de vulnerabilidade e um entrave dependendo do número de requisições realizadas. A proposta original do OpenFlow [McKeown et al. 2008] deixa essa limitação clara já que propõe o uso dessa técnica de comutação em redes de *campi* universitários. Por outro lado, porém, facilita a implementação de algoritmos para escolher quais as ações apropriadas para os fluxos correntes. Uma das vantagens do OpenFlow é o isolamento de redes concorrentes. Nesse caso, mais de um controlador é usado, um para cada rede. O OpenFlow oferece uma camada de abstração denominada FlowVisor [Sherwood et al. 2010] para compartilhamento de recursos físicos entre diferentes redes isoladas.

Dentre as ações que podem ser realizadas pelo OpenFlow estão o encaminhamento de pacotes, remoção de pacotes e a criação de fluxos. Os pacotes de um fluxo podem ser encaminhados para uma determinada porta conforme requisitos desejados. Essa facilidade permite que os pacotes sejam roteados através da rede. Além disso, outra ação relacionada ao encaminhamento de pacotes pode definir que todos os pacotes de um fluxo devem primeiro ser encaminhados para o controlador. Apesar dessa opção reduzir a velocidade de encaminhamento de pacote, ela pode facilitar operações dinâmicas visto que é somente possível executar algoritmos no controlador. Outra ação que pode ser definida pelo OpenFlow é o descarte de alguns ou todos os pacotes relacionados com um determinado fluxo. Já a criação de fluxos pode ocorrer caso haja necessidade de realizar engenharia de tráfego. As ações podem ser usadas para gerenciar o tráfego da rede. É importante notar que todas as ações são definidas pelo controlador. Controladores mais sofisticados como o NOX [Gude et al. 2008] oferecem interfaces para o desenvolvimento de aplicações para gerenciamento ou para obtenção de informações da rede. Dentre possíveis aplicações estão a manutenção de conexões mesmo após o deslocamento de usuários móveis. Nessa aplicação, os pontos de acesso sem-fio monitoram os usuários conectados. Uma vez que um usuário mude de ponto de acesso, o ponto de acesso relata ao controlador que realiza ações para manter as conexões estabelecidas pelo usuário móvel. Essas ações correspondem à mudança do caminho seguido pelos pacotes destinados a esse usuário. Logo, as tabelas de fluxo são ajustadas para que o fluxo correspondente seja encaminhado até o novo ponto de acesso no qual o usuário está conectado.

2.5. Resultados Experimentais

Esta seção apresenta resultados experimentais obtidos com um protótipo realizado para separação de identificação e localização. O objetivo dos experimentos é verificar o desempenho de um dos novos protocolos para mobilidade na Internet a partir de mudanças do ponto de interconexão com a rede cabeada. O protocolo escolhido foi o HIP, descrito na Seção 2.4.1. O cenário dos experimentos, ilustrado pela Figura 2.30(a), é composto de uma estação móvel, um servidor na rede cabeada e dois pontos de acesso sem-fio. A estação móvel é um Laptop IBM Thinkpad Intel Pentium M 1.7 GHz com 512 MB. Já o servidor utilizado é um computador de mesa Intel Core2 Duo 2.4 GHz com 2 GB de memória RAM. Por fim, os pontos de acesso são dois roteadores Linksys WRT54G e WRT350N. A ferramenta utilizada para implementar os protocolos contidos no HIP foi a ferramenta OpenHip versão 0.7 [OpenHIP 2009].

O objetivo do primeiro experimento é verificar o tempo mínimo necessário para o HIP atualizar o endereço IP de uma estação móvel. Para isso, a estação móvel permanece parada e o endereço IP é alterado manualmente duas vezes. O tempo de atualização é, basicamente, o tempo que uma estação sem-fio executando o HIP precisa para notificar as outras estações da rede da sua mudança de endereço. Nesse experimento, a estação móvel permanece parada enviando mensagens ICMP de requisição (`ping`) para um servidor da rede cabeada através do ponto de acesso A. Em um dado momento a estação móvel modifica seu endereço IP para forçar o procedimento de atualização de endereços do HIP. Durante esse período, os pacotes ICMP de resposta enviados para a estação móvel são perdidos já que o endereço IP da estação mudou. A Figura 2.30(b) ilustra o tempo de ida-e-volta (*Round Trip Time* - RTT) das mensagens ICMP enviadas durante 10 segundos de experimento. A figura ilustra os momentos em que as duas alterações de endereço IP foram executadas. Note que após as alterações, o HIP leva, na média, 677 milissegundos para restabelecer a comunicação. Esse tempo é decorrente principalmente da implementação do OpenHip que processa mensagens do HIP apenas periodicamente.



(a) Cenário de teste. (b) Atualização de endereços IP. (c) *Handoff* da estação móvel.

Figura 2.30. Experimentos práticos.

O segundo experimento avalia o impacto da mudança dos endereços IP na transferência de dados utilizando o HIP. Nesse experimento, a estação móvel se desloca do ponto de acesso sem-fio A até o ponto de acesso B, realizando assim um processo de *handoff*. Nesse teste, a estação móvel também envia mensagens de requisição ICMP e espera os pacotes de resposta do servidor na rede cabeada. Durante o processo de *handoff*, a mudança dos pontos de acesso é realizada quando a intensidade do sinal do ponto de acesso B supera a intensidade do sinal do ponto de acesso A. Ao detectar esse evento, a estação móvel conecta-se ao ponto de acesso B e troca de endereço IP. A Figura 2.30(c) mostra que o processo de *handoff* e a troca de mensagens de atualização do HIP duraram 3,354 segundos. Esse tempo é decorrência da troca não instantânea de pontos de acesso, o que resulta em perdas de mensagens de atualização e, consequentemente, maior tempo até o processo de atualização de endereço ser concluído pelo HIP.

Os resultados demonstram que a mobilidade na Internet é viável, mas ainda precisa de ajustes para se tornar transparente aos usuários.

2.6. Considerações Finais

Ao longo dos anos, a Internet vem sofrendo adaptações para atender demandas não previstas originalmente. Tais mudanças são observadas na camada de interconexão

de redes através de novas propostas para substituição e adaptação do IP e dos principais protocolos de roteamento intra e interdomínio. A principal dificuldade dessas propostas é que elas são tentativas de adequar a Internet a uma realidade em constante evolução. A outra possibilidade é o rompimento completo com o projeto inicial da Internet e a adoção de uma nova proposta totalmente inovadora. A experiência adquirida com os anos de operação da Internet poderia ser utilizada em uma proposta evolutiva para a Internet do Futuro. A primeira opção vem sendo adotada pelos provedores de serviço já que implica em mudanças gerenciáveis. Entretanto, por se tratarem de mudanças de curto prazo, até hoje não resolveram por completo os problemas da Internet. Já a segunda opção pode exigir maiores dispêndios financeiros dos provedores de serviço por ser mais radical, em compensação pode ser mais vantajosa no longo prazo. Enquanto esse impasse persiste, demandas emergentes como a mobilidade das estações, os múltiplos domicílios, os múltiplos caminhos e a escalabilidade dos roteadores continuam sem uma solução definitiva satisfatória.

Este minicurso apresentou propostas tanto adaptativas quanto radicais para a interconexão de redes na Internet do Futuro. Algumas propostas se mostraram mais preocupadas com a questão da implementação gradual embora ainda impliquem em mudanças na rede. A lição aprendida é que o problema é complexo e talvez não seja possível atender todas as demandas emergentes e ainda ser escalável e financeiramente factível. Algumas das demandas podem continuar como casos a parte mesmo naquela que será a Internet do Futuro. Uma nova arquitetura que mantenha as mesmas características que fizeram da Internet um dos maiores sucessos do século XX e, ao mesmo tempo, que atenda as novas demandas é um problema em aberto para os próximos anos e ainda despertará grande interesse na comunidade científica.

Agradecimentos

Este trabalho utilizou recursos da CAPES, CNPq, FAPERJ, FUJB, FUNTTEL e FINEP.

Referências

[Alves et al. 2009] Alves, R. S., Campbell, I. V., Couto, R. S., Campista, M. E. M., Moraes, I. M., Rubinstein, M. G., Costa, L. H. M. K., Duarte, O. C. M. B. e Abdalla, M. (2009). *Redes Veiculares: Princípios, Aplicações e Desafios*, capítulo 5, páginas 199–254. Minicursos do Simpósio Brasileiro de Redes de Computadores (SBRC). Sociedade Brasileira de Computação (SBC).

[Andersen et al. 2008] Andersen, D. G., Balakrishnan, H., Feamster, N., Koponen, T., Moon, D. e Shenker, S. (2008). Accountable Internet Protocol (AIP). Em *ACM SIGCOMM*, páginas 339–350.

[Andersen et al. 2001] Andersen, D. G., Balakrishnan, H., Kaashoek, M. F. e Morris, R. (2001). Resilient Overlay Networks. *ACM SIGOPS Operating Systems Review*, 35(5):131–145.

[Ballani e Francis 2007] Ballani, H. e Francis, P. (2007). CONMan: a step towards network manageability. *ACM SIGCOMM Computer Communication Review*,

37(4):205–216.

[Ballani et al. 2008] Ballani, H., Francis, P., Cao, T. e Wang, J. (2008). ViAggre: Making routers last longer! Em *Workshop on Hot Topics in Networks (HotNets-VII)*, páginas 1–6.

[Ballani et al. 2009] Ballani, H., Francis, P., Cao, T. e Wang, J. (2009). Making routers last longer with ViAggre. Em *Symposium on Networked Systems Design and Implementation (NSDI)*, páginas 453–466.

[Brim et al. 2008] Brim, S., Chiappa, N., Farinacci, D., Fuller, V., Lewis, D. e Meyer, D. (2008). LISP-CONS: A Content distribution Overlay Network Service for LISP. IETF Network Working Group Internet-Draft (Trabalho em andamento).

[Caesar et al. 2006a] Caesar, M., Castro, M., Nightingale, E. B., O'Shea, G. e Rowstron, A. (2006a). Virtual Ring Routing: network routing inspired by DHTs. Em *ACM SIGCOMM*, páginas 351–362.

[Caesar et al. 2006b] Caesar, M., Condie, T., Kannan, J., Lakshminarayanan, K., Stoica, I. e Shenker, S. (2006b). ROFL: Routing On Flat Labels. Em *ACM SIGCOMM*, páginas 363–374.

[Caida 2010] Caida (2010). Visualizing IPv4 and IPv6 Internet topology at a macroscopic scale. Acessado 10/03/10 em http://www.caida.org/research/topology/as_core_network/.

[CIDR 2009] CIDR (2009). Aggregation summary. Acessado 18/12/2009 em http://www.cidr-report.org/as2.0/.

[CIDR 2010] CIDR (2010). Active BGP entries (FIB). Acessado 16/03/2010 em http://www.cidr-report.org/as2.0/.

[Clark et al. 2004] Clark, D., Braden, R., Sollins, K., Wroclawski, J., Katabi, D., Kulik, J., Yang, X., Faber, T., Falk, A., Pingali, V., Handley, M. e Chiappa, N. (2004). New Arch: Future generation Internet architecture. Relatório técnico, USC Information Sciences Institute Computer Networks Division, MIT Laboratory for Computer Science and International Computer Science Institute (ICSI).

[Clark et al. 2003] Clark, D. D., Partridge, C., Ramming, J. C. e Wroclawski, J. T. (2003). A knowledge plane for the Internet. Em *ACM SIGCOMM*, páginas 3–10.

[Clausen et al. 2001] Clausen, T., Jacquet, P., Laouiti, A., Muhlethaler, P., Qayyum, A. e Viennot, L. (2001). Optimized link state routing protocol. Em *IEEE International Multi Topic Conference (INMIC)*, páginas 62–68.

[Costa et al. 2006] Costa, L. H. M. K., Fdida, S. e Duarte, O. C. M. B. (2006). Incremental service deployment using the Hop By Hop Multicast Routing Protocol. *IEEE/ACM Transactions on Networking*, 14(3):543–556.

[Crowcroft 2008] Crowcroft, J. (2008). Toward a network architecture that does everything. *Communications of ACM*, 51(1):74–77.

[Deering e Hinden 1998] Deering, S. e Hinden, R. (1998). Internet Protocol, version 6 (IPv6) specification. IETF Network Working Group RFC 2460.

[Deering 1988] Deering, S. E. (1988). Multicast routing in internetworks and extended LANs. Em *ACM SIGCOMM*, páginas 55–64.

[Devarapalli et al. 2005] Devarapalli, V., Wakikawa, R., Petrescu, A. e Thubert, P. (2005). Network Mobility (NEMO) basic support protocol. IETF Network Working Group RFC 3963.

[Farinacci et al. 2009a] Farinacci, D., Fuller, V., Meyer, D. e Lewis, D. (2009a). LISP ALternative Topology (LISP+ALT). IETF Network Working Group Internet-Draft (Trabalho em andamento).

[Farinacci et al. 2009b] Farinacci, D., Fuller, V., Oran, D. e Meyer, D. (2009b). Locator/ID separation protocol (LISP). IETF Network Working Group Internet-Draft (Trabalho em andamento).

[Feldmann et al. 2009] Feldmann, A., Cittadini, L., Mühlbauer, W., Bush, R. e Maennel, O. (2009). HAIR: Hierarchical Architecture for Internet Routing. Em *ACM Workshop on Re-Architecting the Internet (ReArch)*, páginas 43–48.

[Ganesan et al. 2004] Ganesan, P., Gummadi, K. e Garcia-Molina, H. (2004). Canon in G major: Designing DHTs with hierarchical structure. Em *International Conference on Distributed Computing Systems (ICDCS)*, páginas 263–272.

[Godfrey et al. 2009] Godfrey, P. B., Ganichev, I., Shenker, S. e Stoica, I. (2009). Pathlet routing. Em *ACM SIGCOMM*, páginas 111–122.

[Gude et al. 2008] Gude, N., Koponen, T., Pettit, J., Pfaff, B., Casado, M., McKeown, N. e Shenker, S. (2008). NOX: Towards an operating system for networks. *ACM SIGCOMM Computer Communication Review*, 38(3):105–110.

[Hanbali et al. 2005] Hanbali, A. A., Altman, E. e Nain, P. (2005). Survey of TCP over ad hoc networks. *IEEE Communications Surveys & Tutorials*, 7(3):22–36.

[He e Rexford 2008] He, J. e Rexford, J. (2008). Towards Internet-wide multipath routing. *IEEE Network*, 22(2):16–21.

[Hinden e Sheltzer 1982] Hinden, R. e Sheltzer, A. (1982). The DARPA Internet gateway. IETF Network Working Group RFC 823.

[History Museum 2010] History Museum, C. (2010). Internet history. Acessado em http://www.computerhistory.org.

[Iannone e Bonaventure 2007] Iannone, L. e Bonaventure, O. (2007). On the cost of caching locator/ID mappings. Em *ACM CoNEXT*, páginas 1–12.

[Internet World Stats 2010] Internet World Stats (2010). World Internet users and population stats. Acessado em http://www.internetworldstats.com/stats.htm.

[Jen et al. 2009] Jen, D., Meisel, M., Massey, D., Wang, L., Zhang, B. e Zhang, L. (2009). APT: A Practical Tunneling architecture for routing scalability. Relatório técnico, UCLA.

[Jen et al. 2008] Jen, D., Meisel, M., Yan, H., Massey, D., Wang, L., Zhang, B. e Zhang, L. (2008). Towards a new Internet routing architecture: Arguments for separating edges from transit core. Em *ACM Workshop on Hot Topics in Networks (HotNets)*, páginas 1–6.

[Johnson et al. 2004] Johnson, D., Perkins, C. e Arkko, J. (2004). Mobility support in IPv6. IETF Network Working Group RFC 3775.

[Kaur et al. 2003] Kaur, H. T., Kalyanaraman, S., Weiss, A., Kanwar, S. e Gandhi, A. (2003). BANANAS: an evolutionary framework for explicit and multipath routing in the Internet. Em *ACM SIGCOMM workshop on Future Directions in Network Architecture (FDNA)*, páginas 277–288.

[Kim et al. 2009] Kim, C., Caesar, M., Gerber, A. e Rexford, J. (2009). Revisiting route caching: The world should be flat. Em *International Conference on Passive and Active Network Measurement (PAM)*, páginas 3–12.

[Kim et al. 2008] Kim, C., Caesar, M. e Rexford, J. (2008). Floodless in SEATTLE: a Scalable Ethernet Architecture for Large Enterprises. Em *ACM SIGCOMM*, páginas 3–14.

[Krioukov et al. 2007] Krioukov, D., kc claffy, Fall, K. e Brady, A. (2007). On compact routing for the Internet. *ACM SIGCOMM Computer Communication Review*, 37(3):41–52.

[Lear 2010] Lear, E. (2010). NERD: A Not-so-novel EID to RLOC Database. IETF Network Working Group Internet-Draft (Trabalho em andamento).

[Lougheed e Rekhter 1989] Lougheed, K. e Rekhter, Y. (1989). A Border Gateway Protocol (BGP). IETF Network Working Group RFC 1105.

[Luo et al. 2009] Luo, H., Qin, Y. e Zhang, H. (2009). A DHT-based identifier-to-locator mapping approach for a scalable Internet. *IEEE Transactions on Parallel and Distributed Systems*, 20(12):1790–1802.

[Massey et al. 2007] Massey, D., Wang, L., Zhang, B. e Zhang, L. (2007). A scalable routing system design for future Internet. Em *ACM SIGCOMM IPv6 and the Future of the Internet Workshop*, páginas 1–6.

[Mateo 2009] Mateo, M. P. (2009). OpenFlow switching performance. Dissertação de mestrado, Politecnico Di Torino, Torino, Itália.

[McCarthy et al. 2008a] McCarthy, B., Edwards, C. e Dunmore, M. (2008a). Using NEMO to extend the functionality of MANETs. Em *IEEE International Conference on Communications (ICC)*, páginas 455–460.

[McCarthy et al. 2009] McCarthy, B., Edwards, C. e Dunmore, M. (2009). Using NEMO to support the global reachability of MANET nodes. Em *IEEE Conference on Computer Communications (INFOCOM)*, páginas 2097–2105.

[McCarthy et al. 2008b] McCarthy, B., Jakeman, M., Edwards, C. e Thubert, P. (2008b). Protocols to efficiently support nested NEMO (NEMO+). Em *International workshop on Mobility in the evolving Internet Architecture (MobiArch)*, páginas 43–48.

[McKeown et al. 2008] McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S. e Turner, J. (2008). Openflow: Enabling innovation in campus networks. *ACM SIGCOMM Computer Communication Review*, 38(2):69–74.

[Meyer et al. 2007] Meyer, D., Zhang, L. e Fall, K. (2007). Report from the IAB Workshop on Routing and Addressing. Acessado em http://tools.ietf.org/html/draft-iab-raws-report-02.txt.

[Mills 1984] Mills, D. L. (1984). Exterior Gateway Protocol formal specification. IETF Network Working Group RFC 904.

[Moreira et al. 2009] Moreira, M. D. D., Fernandes, N. C., Costa, L. H. M. K. e Duarte, O. C. M. B. (2009). *Internet do Futuro: Um Novo Horizonte*, capítulo 1, páginas 1–59. Minicursos do Simpósio Brasileiro de Redes de Computadores (SBRC). Sociedade Brasileira de Computação (SBC).

[Moskowitz et al. 2008] Moskowitz, R., Nikander, P., Jokela, P. e Henderson, T. (2008). Host Identity Protocol. IETF Network Working Group RFC 5201.

[Motiwala et al. 2008] Motiwala, M., Elmore, M., Feamster, N. e Vempala, S. (2008). Path splicing. Em *ACM SIGCOMM*, páginas 27–38.

[Nikander et al. 2008] Nikander, P., Henderson, T., Vogt, C. e Arkko, J. (2008). End-host mobility and multihoming with the Host Identity Protocol. IETF Network Working Group RFC 5206.

[Nordmark e Bagnulo 2009] Nordmark, E. e Bagnulo, M. (2009). Shim6: Level 3 multihoming Shim protocol for IPv6. IETF Network Working Group RFC 5533.

[OpenHIP 2009] OpenHIP (2009). Host Identity Protocol (HIP). Acessado em http://www.openhip.org/.

[Perera et al. 2004] Perera, E., Sivaraman, V. e Seneviratne, A. (2004). Survey on network mobility support. *ACM SIGMOBILE Mobile Computing and Communications Review*, 8(2):7–19.

[Perkins 2002] Perkins, C. (2002). IP mobility support for IPv4. IETF Network Working Group RFC 3344.

[Ratnasamy et al. 2005] Ratnasamy, S., Shenker, S. e McCanne, S. (2005). Towards an evolvable Internet architecture. *ACM SIGCOMM Computer Communication Review*, 35(4):313–324.

[Rekhter e Li 1995] Rekhter, Y. e Li, T. (1995). A Border Gateway Protocol 4 (BGP-4). IETF Network Working Group RFC 1771.

[Rekhter et al. 2006] Rekhter, Y., Li, T. e Hares, S. (2006). A Border Gateway Protocol 4 (BGP-4). IETF Network Working Group RFC 4271.

[Sabeur et al. 2006] Sabeur, M., Jouaber, B. e Zeghlache, D. (2006). Light-NEMO+: Route optimzation for light-NEMO solution. Em *IEEE International Conference on Networks (ICON)*, páginas 1–6.

[Saucez et al. 2009] Saucez, D., Iannone, L. e Bonaventure, O. (2009). OpenLISP: An open source implementation of the locator/ID separation protocol. Em *ACM SIGCOMM Demos Session*, páginas 1–2.

[Sherwood et al. 2010] Sherwood, R., Chan, M., Covington, A., Gibb, G., Flajslik, M., Handigol, N., Huang, T.-Y., Kazemian, P., Kobayashi, M., Naous, J., Seetharaman, S., Underhill, D., Yabe, T., Yap, K.-K., Yiakoumis, Y., Zeng, H., Appenzeller, G., Johari, R., McKeown, N. e Parulkar, G. (2010). Carving research slices out of your production networks with OpenFlow. *ACM SIGCOMM Computer Communication Review*, 40(1):129–130.

[Stoica et al. 2004] Stoica, I., Adkins, D., Zhuang, S., Shenker, S. e Surana, S. (2004). Internet indirection infrastructure. *IEEE/ACM Transactions on Networking*, 12(2):205–218.

[Stoica et al. 2003] Stoica, I., Morris, R., Liben-Nowell, D., Karger, D., Kaashoek, M., Dabek, F. e Balakrishnan, H. (2003). Chord: A scalable peer-to-peer lookup service for Internet applications. *IEEE/ACM Transactions on Networking*, 11(1):17–32.

[Vogt 2008] Vogt, C. (2008). Six/One Router: A scalable and backwards compatible solution for provider-independent addressing. Em *Workshop on Mobility in the Evolving Internet Architecture (MobiArch)*, páginas 13–18.

[Wetherall 2006] Wetherall, X. Y. D. (2006). Source selectable path diversity via routing deflections. Em *ACM SIGCOMM*, páginas 159–170.

[Xu e Rexford 2006] Xu, W. e Rexford, J. (2006). MIRO: Multi-path Interdomain ROuting. Em *ACM SIGCOMM*, páginas 171–182.

[Yang 2003] Yang, X. (2003). NIRA: a New Internet Routing Architecture. Em *ACM SIGCOMM workshop on Future Directions in Network Architecture (FDNA)*, páginas 301–312.

[Yang et al. 2007] Yang, X., Clark, D. e Berger, A. W. (2007). NIRA: a New Interdomain Routing Architecture. *IEEE/ACM Transactions on Networking*, 15(4):775–788.