3
Projeto de rede

Antes de adquirir algum equipamento ou decidir qual plataforma de hardware utilizar, você deve ter uma idéia clara da natureza de seu problema de comunicação. Provavelmente você está lendo este livro porque necessita conectar redes de computadores de forma a compartilhar recursos e, afinal, conectar-se à Internet global. O projeto de rede que você decidir implementar deve estar adequado ao problema de comunicação que você deseja resolver. Você precisa conectar um local remoto à uma ligação com a Internet no centro de seu campus? Sua rede tem a possibilidade de crescer para incorporar vários locais remotos? A maior parte dos componentes de sua rede será instalada em locais fixos, ou sua rede irá expandir para incluir centenas de laptops e outros dispositivos móveis?

Neste capítulo vamos revisar os conceitos de rede que definem o TCP/IP, a família principal de protocolos de rede utilizada na Internet. Veremos exemplos de como outras pessoas construíram suas redes wireless para resolver seus problemas de comunicação, incluindo os diagramas da estrutura essencial da rede. Finalmente, apresentaremos vários métodos comuns para garantir o fluxo eficiente da informação dentro de sua rede e também para o restante do mundo.

Entendendo redes

O nome TPC/IP refere-se a um conjunto de protocolos que permite a troca de informações na Internet global. Com o conhecimento do TCP/IP, você poderá construir redes que poderão crescer para, virtualmente, qualquer tamanho e, objetivamente, permitir que sua rede faça parte da Internet.

Caso você sinta-se confortável com os fundamentos de redes com TCP/IP (incluindo endereçamento, roteamento, switches, firewalls e roteadores), você pode ir direto para **Projetando a rede física**, na **Página 51**. Agora iremos rever os conceitos básicos de rede Internet.

Introdução

Veneza, na Itália, é uma cidade fantástica para nos perdermos. As estradas são meros caminhos que cruzam a água em centenas de lugares, nunca em

uma linha reta. Os funcionários de correio em Veneza estão entre os mais altamente treinados do mundo, cada um especializado na entrega de correspondência para apenas um ou dois dos seis *sestieri* (distritos) da cidade. Isto acontece em função do intrincado projeto dessa cidade ancestral. Muitas pessoas descobrem que saber a localização da água e do sol é muito mais útil do que tentar encontrar o nome de uma rua em um mapa.

Figura 3.1: Outro tipo de máscara de rede.

Imagine um turista que acaba de encontrar, como um suvenir, uma máscara de papel-machê, e quer enviá-la do estúdio em S. Polo, Veneza, para um escritório em Seattle, nos Estados Unidos. Isto pode parecer uma tarefa ordinária, mesmo trivial, mas vamos observar o que realmente acontece.

O artista, primeiramente, coloca a máscara em um pacote apropriado para a remessa e a endereça para o escritório em Seattle, nos Estados Unidos. Depois, a entrega a um empregado do correio, que anexa alguns formulários oficiais ao pacote e o envia para um departamento central de processamento, que lida com a remessa para localidades internacionais. Depois de vários dias, o pacote é liberado pela alfândega italiana e segue seu destino em um vôo transatlântico, chegando a uma central de processamento de importações nos Estados Unidos. Uma vez que é liberado pela alfândega americana, o pacote é enviado pra um ponto de distribuição no noroeste dos Estados Unidos e, daí, para a central de processamento dos correios de Seattle. Então, o pacote segue seu caminho em um carro de entregas que cumpre uma rota que o leva para um determinado endereço, em uma determinada rua, dentro de um determinado bairro. Um recepcionista, no escritório, aceita o pacote e o coloca na caixa de correio apropriada. Finalmente, o pacote é retirado pelo destinatário e a máscara chega a seu destino.

O recepcionista no escritório de Seattle não sabe ou sequer se importa em como se faz para chegar no sestiere de S. Polo, em Veneza. Seu trabalho é simplesmente receber os pacotes que chegam e colocá-los na caixa de correio da pessoa que irá recebê-los. De forma similar, o carteiro em Veneza não precisa preocupar-se em como chegar ao bairro correto em Seattle. Seu trabalho é pegar os pacotes de sua vizinhança local e encaminhá-los ao posto de correio mais próximo na cadeia de remessas.

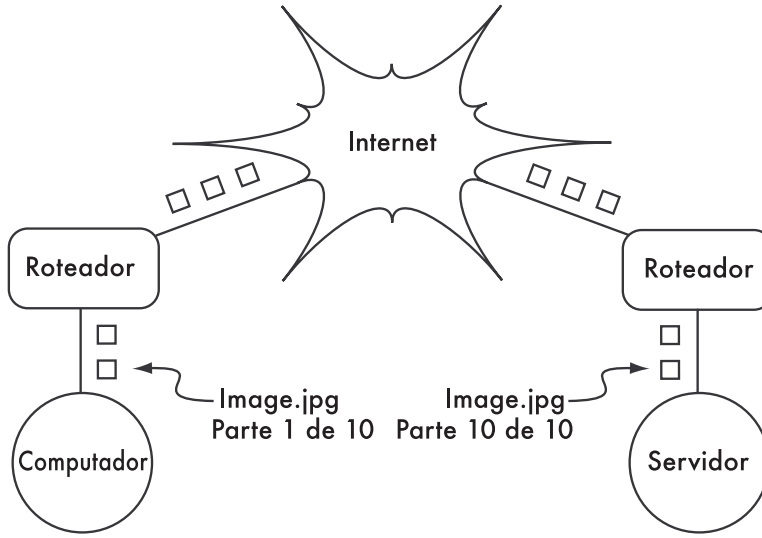


Figura 3.2: Transmissão de rede internet. Pacotes são encaminhados entre os roteadores até que atinjam seu destino final.

Isto é bastante similar à forma pela qual o roteamento na Internet funciona. Uma mensagem é dividida em uma série de ***pacotes*** individuais, que são etiquetados com seu remetente (fonte) e destinatário (ou destino). O computador envia, então, estes pacotes para um ***roteador***, que decide para onde enviá-los a seguir. O roteador apenas precisa conhecer um punhado de rotas (por exemplo, como chegar à rede local, a melhor rota para algumas outras redes locais e uma rota para o caminho de acesso—*gateway*—que conecta ao resto da Internet). Esta lista de possíveis rotas é chamada de ***tabela de roteamento*** (em inglês, *routing table*). Na medida em que os pacotes chegam ao roteador, o endereço de destino é examinado e comparado com a sua tabela de roteamento interna. Caso o roteador não tenha nenhuma rota explícita para o destino em questão, ele envia o pacote para o destino mais parecido que consiga encontrar, que é, tipicamente, o próprio gateway para a Internet (através da ***rota padrão***, ou *default route*). O próximo roteador executa o mesmo processo, e assim sucessivamente, até que o pacote chegue a seu destino.

Pacotes apenas podem seguir seu caminho pelo sistema postal internacional porque estabelecemos um esquema de endereçamento padrão para eles. Por exemplo, o endereço destino deve estar escrito de forma legível na frente do pacote e incluir toda a informação crítica (como o nome do destinatário, rua e número, cidade, estado, país e código de endereçamento postal). Sem esta informação, os pacotes são devolvidos para o remetente ou perdem-se pelo caminho.

Pacotes apenas podem trafegar pela Internet global porque foi possível chegar a um acordo sobre um esquema de endereçamento e protocolo para o encaminhamento dos mesmos. Estes protocolos de comunicação padrão permitem a troca de informação em uma escala global.

Comunicações cooperativas

A comunicação só é possível quando os participantes falam em um idioma comum. Mas quando a comunicação torna-se mais complexa que uma simples conversa entre duas pessoas, o protocolo torna-se tão importante quanto o idioma. Todas as pessoas em um auditório podem falar inglês, mas sem um conjunto de regras que estabeleçam quem tem o direito de uso do microfone, a comunicação das idéias de um indivíduo para todos os demais será praticamente impossível. Agora, imagine um auditório tão grande quanto o mundo, cheio de todos os computadores existentes. Sem um conjunto comum de protocolos de comunicação que regule quando e como cada computador pode falar, a Internet seria uma bagunça caótica onde todas as máquinas tentam falar ao mesmo tempo.

Uma série de modelos de comunicação foi desenvolvida para resolver este problema. O mais conhecido é o ***modelo OSI***.

O modelo OSI

O padrão internacional para a interconexão de sistemas abertos (OSI—Open Systems Interconection) está definido no documento ISO/IEC 7498-1, referendado pela International Standards Organization e pela International Electrotechnical Commission. O padrão completo está disponível na publicação "ISO/IEC 7498-1:1994," que pode ser encontrada em *http://standards.iso.org/ittf/PubliclyAvailableStandards/*.

O modelo OSI divide o tráfego de rede em um número de ***camadas***. Cada camada é independente das demais camadas ao redor dela e cada uma constrói, a partir dos serviços entregues pela camada inferior, novos serviços que provê para a camada superior. A abstração entre as camadas torna fácil o projeto de elaboradas e altamente confiáveis ***pilhas de protocolos*** (*protocol stacks*), como a onipresente pilha ***TCP/IP***. Uma pilha de protocolo é uma real implementação de um ambiente de rede em camadas. O modelo OSI não define qual o protocolo a ser utilizado em uma determinada rede, mas simplesmente delega quais trabalhos de comunicação são executados por cada camada, dentro de uma hierarquia bem definida.

Enquanto a especificação ISO/IEC 7498-1 detalha como as camadas devem interagir uma com as outras, ela deixa os detalhes desta implementação para os fabricantes. Cada camada pode ser implementada em hardware (comumente, as mais baixas) ou em software. Desde que a interface entre as camadas esteja de acordo com o padrão, os fabricantes estão à vontade para usar qualquer recurso disponível para construir sua pilha de protocolo. Isto significa que qualquer camada de um fabricante A pode interoperar com a mesma camada do fabricante B (desde que as especificações relevantes tenham sido interpretadas e implementadas corretamente).

Aqui está uma breve descrição das sete camadas do modelo de rede OSI:

Camada	Nome	Descrição
7	Aplicação	A ***Camada de Aplicação*** é a camada à qual a maior parte dos usuários da rede estão expostos e é o nível no qual a comunicação humana acontece. HTTP, FTP e SMTP são todos protocolos da camada de aplicação. O humano fica acima desta camada, interagindo com a aplicação.
6	Apresentação	A ***Camada de Apresentação*** é responsável por lidar com a representação dos dados, antes que eles cheguem à aplicação. Isto pode incluir a codificação MIME, compressão de dados, verificação de formatos, ordenação de bytes, etc.
5	Sessão	A ***Camada de Sessão*** gerencia as sessões de comunicação lógica entre aplicações. NetBIOS e RPC são dois exemplos de protocolo da camada cinco.
4	Transporte	A ***Camada de Transporte*** fornece um método para o acesso a um serviço específico em um dado nó da rede. Exemplos de protocolos que operam nesta camada são o TCP e o UDP. Alguns protocolos na camada de transporte (como o TCP) garantem que todos os dados cheguem ao destino, sejam rearranjados e entregues à próxima camada na ordem correta. UDP é um protocolo "*connectionless*" (sem conexão) comumente usado para a transmissão (*streaming*) de áudio e vídeo.
3	Rede	IP (o Protocolo de Internet, *Internet Protocol*) é o mais comum protocolo de ***Camada de Rede***. Esta é a camada onde o roteamento ocorre. Pacotes podem deixar a conexão de rede local e ser retransmitidos para outras redes. Roteadores executam esta função na rede tendo ao menos duas interfaces de rede, uma para cada rede que interconectam. Nós na Internet são alcançados pelo seu único, globalmente individual, endereço IP. Outro protocolo de rede crítico é o ICMP, um protocolo especial que fornece várias mensagens de gerenciamento necessárias para a correta operação do IP. Esta camada é chamada também, algumas vezes, de ***Camada de Internet***.

Camada	Nome	Descrição
2	Conexão de Dados	Sempre que dois ou mais nós compartilham o mesmo meio físico (por exemplo, vários computadores conectados em um hub, ou uma sala cheia de dispositivos wireless, todos usando o mesmo canal de rádio), eles utilizam a ***Camada de Conexão de Dados*** para comunicarem-se. Exemplos de protocolos de conexão de dados são Ethernet, Token Ring, ATM e os vários protocolos wireless (802.11 a/b/g). A comunicação nesta camada é dita link-local, uma vez que todos os nós conectados nesta camada comunicam-se, um com o outro, diretamente. Esta camada também é conhecida como ***Media Access Contro***l (***MAC*** — Controle de Acesso ao Meio). Em redes modeladas com base na Ethernet, os nós são referenciados por seu ***endereço MAC***. Ele é composto de um número de 48 bits designado de forma única e individual para cada dispositivo de rede quando o mesmo é fabricado.
1	Física	A ***Camada Física*** é a mais baixa camada no modelo OSI e refere-se ao próprio meio físico no qual a comunicação ocorre. Ela pode ser um cabo de cobre de categoria 5 (***CAT 5***), um feixe de fibra ótica, ondas de rádio ou qualquer outro meio capaz de transmitir sinais. Cabos cortados, fibras quebradas e interferência de RF são exemplos de problemas da camada física.

As camadas neste modelo são numeradas de um a sete, com o sete no topo. Isto é feito para reforçar a idéia de que cada camada constrói-se acima, e depende da camada abaixo. Imagine o modelo OSI como um prédio, onde a fundação é a primeira camada, as camadas seguintes são os sucessivos andares e o telhado é a camada sete. Caso você remova qualquer uma das camadas, o prédio não se sustenta. De maneira similar, se o quarto andar está em chamas, ninguém consegue passar acima, ou abaixo dele.

As primeiras três camadas (Física, Conexão de Dados e Rede) acontecem todas "na rede". Isto quer dizer que a atividade nestas três camadas é determinada pela configuração de cabos, switches, roteadores e dispositivos similares. Um switch de rede somente pode distribuir pacotes utilizando endereços MAC, assim, ele necessita implementar apenas as camadas um e dois. Um roteador simples pode rotear pacotes utilizando apenas seus endereços IP, então ele necessita implementar as camadas um a três. Um servidor web ou um computador laptop executam aplicações, então eles devem implementar todas as sete camadas. Alguns roteadores avançados podem implementar a camada quatro e acima, permitindo que eles tomem decisões baseadas no conteúdo de alto nível de um pacote, como o nome de um website ou os anexos de um email.

O modelo OSI é reconhecido internacionalmente e é amplamente considerado como o completo e definitivo modelo de rede. Ele fornece aos fabricantes e implementadores de protocolos um ambiente para a construção de dispositivos de rede que interoperam em qualquer parte do mundo.

Pela perspectiva de um engenheiro ou analista de redes, o modelo OSI pode parecer desnecessariamente complexo. Em particular, pessoas que fazem a implementação e a análise de redes TCP/IP raramente necessitam lidar com problemas das camadas de Sessão ou Apresentação. Para a maioria das implementações de conexões de rede Internet, o modelo OSI pode ser simplificado em uma coleção menor, de cinco camadas.

O modelo TCP/IP

De maneira diversa do modelo OSI, o modelo TCP/IP não é um padrão internacional e suas definições variam. Mesmo assim, ele é freqüentemente utilizado como um modelo prático para o entendimento e diagnóstico de redes Internet. A absoluta maioria da Internet usa TCP/IP e, assim, podemos fazer algumas suposições sobre redes que as tornarão mais fáceis de entender. O modelo TCP/IP para redes descreve as quatro camadas a seguir:

Camada	Nome
5	Aplicação
4	Transporte
3	Internet
2	Conexão (link) de Dados
1	Física

Nos termos do modelo OSI, as camadas cinco a sete ficam reunidas, aqui, na camada superior (a Camada de Aplicação). As primeiras quatro camadas, em ambos os modelos, são idênticas. Muitos engenheiros de rede pensam nas camadas acima da quatro como "apenas dados" que variam de aplicação para aplicação. Uma vez que as primeiras três camadas são interoperáveis entre os equipamentos de praticamente qualquer fornecedor, a camada quatro funciona entre quaisquer computadores rodando TCP/IP e o que está acima desta camada tende a relacionar-se a aplicações específicas. Este modelo simplificado funciona bem para a construção e diagnóstico de redes TCP/IP. Usaremos o modelo TCP/IP sempre que discutirmos redes neste livro.

O modelo TCP/IP pode ser comparado com uma pessoa entregando uma carta em um edifício de escritórios no centro da cidade. Esta pessoa primeiro precisa interagir com as próprias ruas (a camada Física), prestar atenção ao tráfego nessas ruas (a camada de Conexão de Dados), virar em uma ou outra rua para chegar ao endereço correto (a camada de Internet), dirigir-se ao prédio, ao andar e ao escritório correto (a camada de Transporte) e, finalmente, entregar

a carta ao recepcionista do escritório que se encarregará dela daí em diante (a camada de Aplicação). Uma vez que a carta foi entregue ao recepcionista, a pessoa que a entregou está livre para seguir seu caminho.

As cinco camadas podem ser facilmente lembradas usando o mnemônico, em inglês, "**P**lease **D**on't **L**ook **I**n **T**he **A**ttic," que corresponde a "**P**hysical / **D**ata **L**ink / **I**nternet / **T**ransport / **A**pplication."

Uma sugestão de um mnemônico em português é: "**F**aça **C**omo **D**efinido, **I**gnorando **T**rabalhos **A**lternativos", correspondendo às camadas "**F**ísica / **C**onexão de **D**ados / **I**nternet / **T**ransporte e **A**plicação".

Os protocolos Internet

TCP/IP é a pilha de protocolos mais comumente utilizada na Internet global. A sigla significa ***Transmission Control Protocol*** (***TCP***—Protocolo de Controle de Transmissão) e ***Internet Protocol*** (***IP***—Protocolo de Internet), mas de fato refere-se a toda uma família de protocolos de comunicação relacionados. O TCP/IP é também conhecido como a ***suíte de protocolo Internet*** (*Internet Protocol suite*), e opera nas camadas três e quatro do modelo TCP/IP.

Nesta discussão, manteremos o foco na versão quatro do protocolo IP (IPv4), uma vez que este é o protocolo mais utilizado na Internet.

Endereçamento IP

Em uma rede IPv4, o endereço é um número de 32 bits, normalmente escrito como quatro números de oito bits expressos no formato decimal e separados por pontos. Exemplos de endereços IP são 10.0.17.1, 192.168.1.1 ou 172.16.5.23.

Se você escrever todos os endereços IP possíveis, eles irão variar de 0.0.0.0 a 255.255.255.255. Isto leva a um total de mais de quatro bilhões de possíveis endereços IP (255 x 255 x 255 x 255 = 4.228.250.625), ainda que alguns deles estejam reservados para funções especiais e não sejam designados a servidores. Cada um dos endereços IP utilizáveis é um identificador único, que distingue um nó da rede de outro.

Redes interconectadas devem chegar a um acordo em um plano de endereçamento IP. Cada endereço IP deve ser único e, geralmente, não poderá ser usado em diferentes lugares na Internet ao mesmo tempo, de outra forma, os roteadores não saberiam a qual deveriam rotear seus pacotes.

Endereços IP são alocados por uma autoridade central de endereçamento, que fornece um método consistente e coerente de numeração. Isto garante que endereços duplicados não serão usados por diferentes redes. Esta autoridade designa grandes blocos de endereços consecutivos a autoridades menores, que por sua vez designam blocos menores destes endereços consecutivos para outras autoridades ou para seus clientes. Estes grupos de endereços são chamados de sub-redes, ou ***subnets***. Grandes sub redes podem ser posteriormente divididas em sub-redes menores. Um grupo de endereços relativos é chamado de um espaço de endereçamento, ou ***address space***.

Subredes

Aplicando uma máscara de sub-rede (***subnet mask***, também chamada de ***network mask*** ou simplesmente ***netmask***) a um endereço IP, você consegue definir logicamente tanto o servidor (host) quanto a rede a qual ele pertence. Tradicionalmente, máscaras de sub-rede são expressas utilizando-se o formato decimal, de forma similar a um endereço IP. Por exemplo, 255.255.255.0 é uma máscara de rede comum. Você irá se deparar com este tipo de notação quando estiver configurando interfaces de rede, criando rotas, etc. Entretanto, máscaras de sub redes são expressas de forma mais sucinta utilizando a ***notação CIDR***, que simplesmente enumera o número de bits em uma máscara após uma barra (/). Assim, 255.255.255.0 pode ser simplificada como /24. CIDR é a abreviatura de ***Classless Inter-Domain Routing*** (roteamento inter-domínios sem classe), e está definido no RFC1518[1].

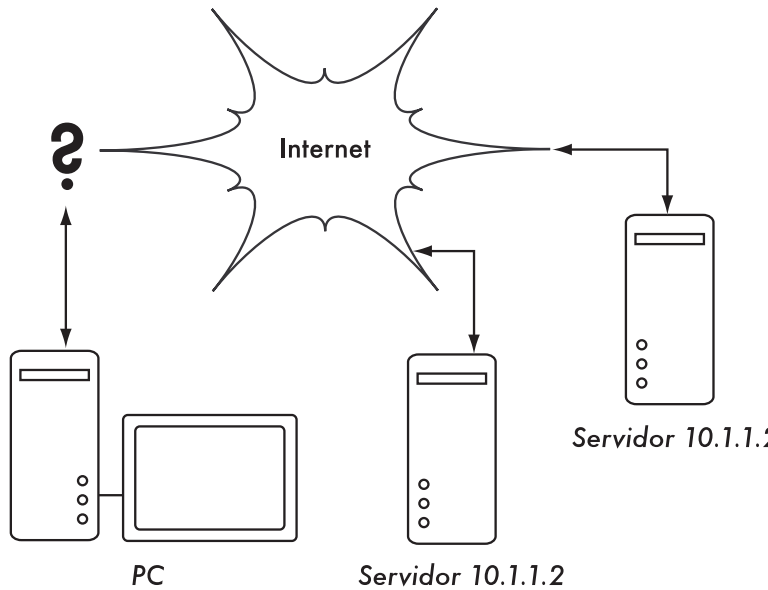


Figura 3.3: Sem que exista um único endereço IP, a não ambigüidade global de roteamento é impossível. Se um PC solicita uma página web do servidor 10.1.1.2, em qual servidor chegará?

Uma máscara de sub-rede define o tamanho de uma determinada rede. Por exemplo, em uma máscara do tipo /24, 8 bits são reservados para servidores (32 bits no total—24 bits da máscara de rede = 8 bits para hosts). Isto permite um total de 256 endereços de servidores (2^8=256). Por convenção, o primeiro valor é reservado para o ***endereço da rede*** (*network address*, .0 ou 00000000), e o

1. RFC é a abreviatura de *Request For Comments* (Solicitação de Comentários). Os RFCs são uma série numerada de documentos publicados pela Internet Society, que registram idéias e conceitos relacionados a tecnologias de Internet. Nem todos os RFCs são, de fato, padrões. Os RFCs estão disponíveis online em *http://rfc.net/*

último valor é definido como o ***endereço de broadcast*** (.255 ou 11111111). Isto deixa 254 endereços disponíveis para os servidores desta rede.

Máscaras de sub-rede trabalham com a aplicação da operação lógica AND (E) a um número IP de 32 bits. Em notação binária, os bits "1" na máscara indicam a porção do endereço de rede, e os bits "0" indicam a porção de endereço do servidor. Uma operação lógica AND é feita comparando dois bits. O resultado é "1" se os dois bits comparados são, ambos, "1". Caso contrário, o resultado é "0". Aqui estão todos os possíveis resultados de uma comparação com a operação AND entre dois bits.

Bit 1	Bit 2	Resultado
0	0	0
0	1	0
1	0	0
1	1	1

Para entender como uma máscara de sub-rede é aplicada a um endereço IP, primeiro converta todos os números para seu equivalente em base binária. A máscara 255.255.255.0 em binário contém 24 bits "1".

```
     255.     255.     255.       0
11111111.11111111.11111111.00000000
```

Quando esta máscara é combinada com um endereço IP 10.10.10.10, podemos aplicar uma operação lógica AND para cada um dos bits, a fim de determinar o endereço da rede.

```
  10.10.10.10: 00001010.00001010.00001010.00001010
255.255.255.0: 11111111.11111111.11111111.00000000
        -----------------------------------
 10.10.10.0: 00001010.00001010.00001010.00000000
```

Isto resulta na rede 10.10.10.0/24. Esta rede consiste de servidores cujo endereço IP varia de 10.10.10.1 até 10.10.10.254, com 10.10.10.0 como o endereço da rede e 10.10.10.255 como o endereço de broadcast.

Máscaras de sub-rede não estão limitadas a octetos inteiros. É possível especificar máscaras como 255.254.0.0 (ou /15 CIDR). Este é um grande bloco, contendo 131.072 endereços, de 10.0.0.0 até 10.1.255.255. Ele pode ser posteriormente subdividido, por exemplo, em 512 sub-redes de 256 endereços cada. A primeira teria os endereços entre 10.0.0.0 e 10.0.0.255, a seguinte entre 10.0.1.0 e 10.0.1.255, e assim sucessivamente até chegar a 10.1.255.0 e 10.1.255.255. Alternativamente, ela poderia ser dividida em dois blocos de 65.536 endereços, ou 8192 blocos de 16 endereços, ou ainda de muitas outras maneiras. Ela poderia até ser dividida em uma mescla de blocos de endereços

de diferentes tamanhos, desde que eles não tenham intersecções entre si e cada um tenha uma sub-rede válida, com o número de endereços representado por uma potência de dois.

Mesmo que muitas máscaras de rede sejam possíveis, os tipos comuns incluem:

CIDR	Decimal	Número de Hosts
/30	255.255.255.252	4
/29	255.255.255.248	8
/28	255.255.255.240	16
/27	255.255.255.224	32
/26	255.255.255.192	64
/25	255.255.255.128	128
/24	255.255.255.0	256
/16	255.255.0.0	65.536
/8	255.0.0.0	16.777.216

A cada redução do valor CIDR, o espaço de endereço IP é dobrado. Lembre-se que dois endereços IP dentro de cada rede sempre estão reservados para o endereço da rede e o de broadcast.

Três máscaras comuns têm nomenclatura especial. Uma rede /8 (com uma máscara de rede de 255.0.0.0) define uma rede ***Classe A***. Uma /16 (255.255.0.0) é uma ***Classe B*** e uma /24 (255.255.255.0) é chamada de ***Classe C***. Estes nomes já eram usados muito antes da notação CIDR, mas seguem em uso por questões históricas.

Endereçamento IP global

Você já se perguntou quem controla a distribuição de endereços IP? ***Endereços IP roteáveis globalmente*** (*Globally routable IP addresses*) são atribuídos e distribuídos por ***Registradores Regionais de Internet*** (*Regional Internet Registrars*—***RIR***s) para provedores de acesso à Internet (*Internet Services Providers*—***ISP***s). Um ISP irá alocar blocos menores de endereços IP para seus clientes quando solicitados. Praticamente todos os usuários da Internet obtêm seus endereços IP de um ISP.

Os quatro bilhões de endereços IP disponíveis são administrados pela ***Internet Assigned Numbers Authority*** (***IANA***, *http://www.iana.org/*). A IANA dividiu este espaço em grandes subnets, usualmente do tipo /8, com 16 milhões de endereços cada. Estas subnets são delegadas para uma das cinco registradoras regionais de Internet (RIRs), para as quais é dada a autoridade sobre grandes áreas geográficas.

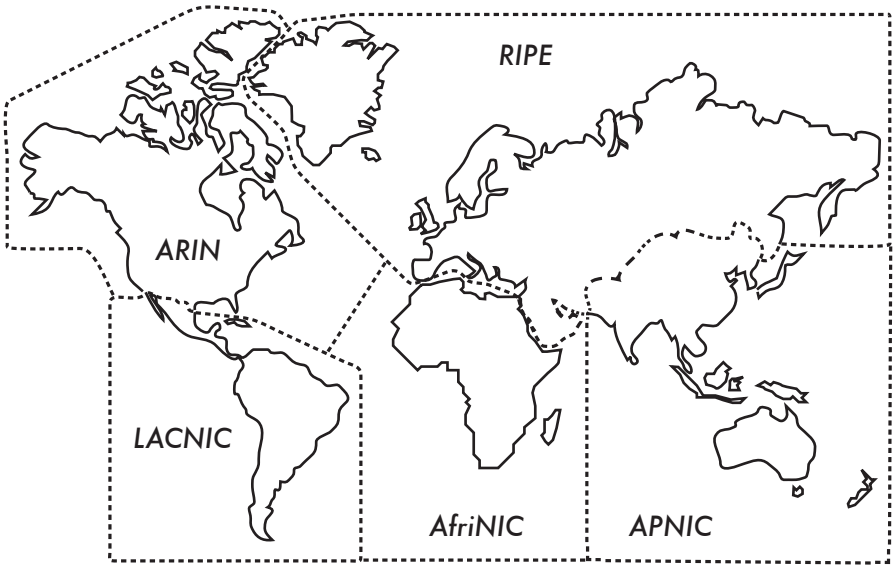


Figura 3.4: Autoridade para a designação de endereços IP é delegada aos cinco Regional Internet Registrars.

Os cinco RIRs são:

- African Network Information Centre (AfriNIC, *http://www.afrinic.net/*)
- Asia Pacific Network Information Centre (APNIC*, http://www.apnic.net/*)
- American Registry for Internet Numbers (ARIN, *http://www.arin.net/*)
- Regional Latin-American and Caribbean IP Address Registry (LACNIC, *http://www.lacnic.net/*)
- Réseaux IP Européens (RIPE NCC, *http://www.ripe.net/*)

Seu provedor de acesso irá fornecer-lhe um espaço de endereços IP roteáveis globalmente, que vem do conjunto de endereços alocado a ele pelo RIR correspondente. O sistema de registro garante que os endereços IP não serão reutilizados em nenhum local da rede, em todo o mundo.

Uma vez que há uma concordância na destinação de endereços IP, é possível transportar pacotes entre as redes e participar da Internet global. Este processo de mover pacotes entre redes é chamado de ***roteamento***.

Endereços IP estáticos

Um endereço IP estático é uma alocação de endereço que nunca muda. Endereços IP estáticos são importantes pois os servidores que usam estes endereços podem ter mapeamentos de DNS apontados para eles e, tipicamente,

eles fornecem informações para outras máquinas (como serviços de email, servidores web, etc.).

Blocos de endereços IP estáticos podem ser disponibilizados por seu provedor de acesso, tanto mediante uma requisição ou automaticamente, dependendo da maneira pela qual você conecta-se à Internet.

Endereços IP Dinâmicos

Endereços IP dinâmicos são atribuídos por um provedor de acesso à Internet para conexões que não são nós permanentes para a Internet, como um computador doméstico que utiliza uma conexão discada.

Endereços IP dinâmicos podem ser atribuídos automaticamente através do protocolo de configuração dinâmica do servidor, o ***Dynamic Host Configuration Protocol*** (***DHCP***) ou do protocolo ponto-a-ponto, ***Point-to-Point Protocol*** (***PPP***), dependendo do tipo de conexão à Internet. Um nó que utiliza DHCP primeiramente solicita um endereço IP da rede e, automaticamente, configura sua interface para o acesso. Endereços IP podem ser atribuídos de forma aleatória, a partir de um conjunto de endereços possíveis, pelo provedor de acesso, ou podem ser atribuídos de acordo com determinada política. Endereços IP atribuídos por DHCP são válidos por um período de tempo (chamado de ***lease time***, tempo de "empréstimo"). O nó deve renovar seu "empréstimo" de DHCP antes que o lease time expire. No momento da renovação, o nó pode receber o mesmo endereço IP ou outro diferente, vindo do conjunto de endereços disponíveis.

Endereços dinâmicos são populares entre os provedores de acesso à Internet, pois esta técnica permite que eles usem endereços IP em menor quantidade do que o número de seus clientes. Eles precisam de um endereço para cada ***cliente ativo em um dado momento***. Endereços IP globalmente roteáveis custam dinheiro, e algumas autoridades que são especializadas na atribuição de endereços (como a RIPE, a RIR da Europa) são bastante estritos na entrega de endereços de IP para provedores de acesso. A atribuição dinâmica de endereços permite economia aos provedores de acesso, e eles tipicamente irão cobrar um valor adicional para oferecer um endereço IP estático a seus clientes.

Endereço IP privado

Muitas redes privadas, em empresas, não necessitam alocar, para todos os seus computadores, endereços IP que possam ser globalmente roteados ou que estejam públicos para a Internet. Particularmente, computadores que não são servidores públicos não precisam estar visíveis na Internet. As empresas normalmente utilizam endereços IP do espaço de endereços privados (***private address space***) para máquinas de sua rede interna.

Existem, atualmente, três blocos de espaços de endereços privados reservados pela IANA: 10.0.0.0/8, 172.16.0.0/12 e 192.168.0.0/16. Eles estão definidos no RFC1918. Estes endereços não são destinados ao roteamento na Internet e são, tipicamente, únicos apenas dentro de uma organização ou agrupamento de organizações que escolherem utilizar um mesmo esquema de endereçamento.

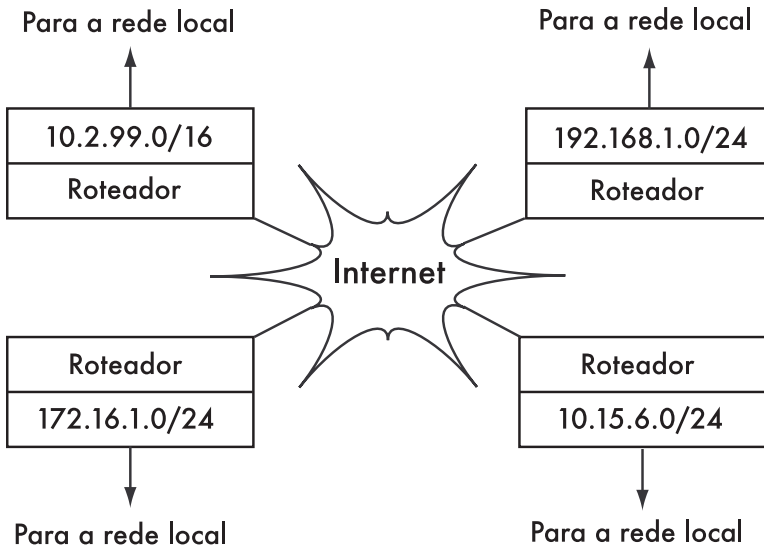


Figura 3.5: Conforme o RFC1918, endereços privados podem ser usados dentro de uma organização e não serão roteados na Internet global.

Caso você pretenda conectar redes privadas que usam o espaço de endereços do RFC1918, certifique-se de que estará usando endereços únicos em todas as redes em questão. Por exemplo, você pode dividir o espaço de endereços 10.0.0.0/8 em múltiplas redes Classe B (10.1.0.0/16, 10.2.0.0/16, etc.). Um bloco pode ser designado para cada rede, de acordo com sua localização física (a sede do campus da Universidade, o primeiro escritório remoto, o segundo escritório remoto, casa do estudante e assim por diante). Os administradores de rede em cada localização podem ainda subdividir a rede em múltiplas sub-redes Classe C (10.1.1.0/24, 10.1.2.0/24, etc.) ou em blocos de qualquer outro tamanho lógico. Assim, caso todas estas redes venham a ser conectadas no futuro (seja por uma conexão cabeada, wireless ou VPN), todas as máquinas poderão ser acessadas de qualquer ponto da rede sem a necessidade de rearranjo de endereços.

Alguns provedores de acesso podem alocar endereços privados como os descritos, ao invés de endereços públicos, a seus clientes, ainda que isto traga sérias desvantagens. Uma vez que estes endereços não podem ser roteados na Internet, os computadores que os usam não são realmente "parte" da Internet, não podendo ser acessados através dela. Para que eles possam comunicar-se com a Internet, seus endereços privados devem ser traduzidos para endereços públicos. Este processo de tradução é conhecido por Tradução de Endereço de Rede (Network Address Translation—NAT), e é normalmente feito pelo gateway entre a rede privada e a Internet. Estudaremos o NAT com mais detalhe na **Página 44**.

Roteamento

Imagine uma rede com três servidores: A, B e C. Eles usam os respectivos endereços IP: 192.168.1.1, 192.168.1.2 e 192.168.1.3. Estes servidores são parte de uma rede /24 (sua máscara é 255.255.255.0).

Para que dois servidores comuniquem-se em uma rede local, eles devem conhecer o endereço MAC um do outro. É possível configurar manualmente cada servidor com uma tabela de mapeamento entre endereços IP e MAC, mas o protocolo de resolução de endereço (***Address Resolution Protocol—ARP***) é normalmente usado para fazer automaticamente esta tarefa.

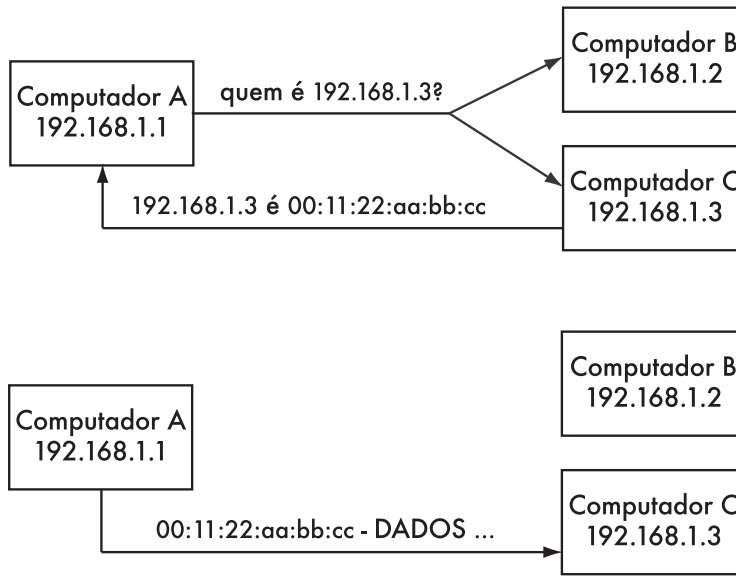


Figura 3.6: O computador A necessita enviar dados para o endereço 192.168.1.3. Mas antes disto, ele deve primeiro perguntar a toda a rede qual é o MAC address que responde pelo endereço 192.168.1.3.

Através da utilização do ARP, o computador A envia uma mensagem geral (broadcast) para todos os demais computadores: "Quem possui o endereço MAC para o IP 192.168.1.3?". Quando o computador C vê a solicitação ARP para o seu próprio endereço IP, ele responde com o seu endereço MAC.

Considere agora outra rede com três hosts: D, E e F, com os respectivos endereços IP 192.168.2.1, 192.168.2.2 e 192.168.2.3. Esta é mais uma rede /24, mas em um espaço de endereços diferente da rede acima. Todos os três hosts conseguem comunicar-se diretamente (primeiro utilizando o ARP para traduzir os endereços IP em endereços MAC, e depois enviando pacotes para o endereço MAC).

Agora vamos adicionar o host G. Este computador tem dois cartões de rede, cada um conectado a uma rede. O primeiro cartão de rede usa o endereço 192.168.1.4 e o outro o 192.168.2.4. O host G é agora um link local para ambas as redes, e pode rotear pacotes entre elas.

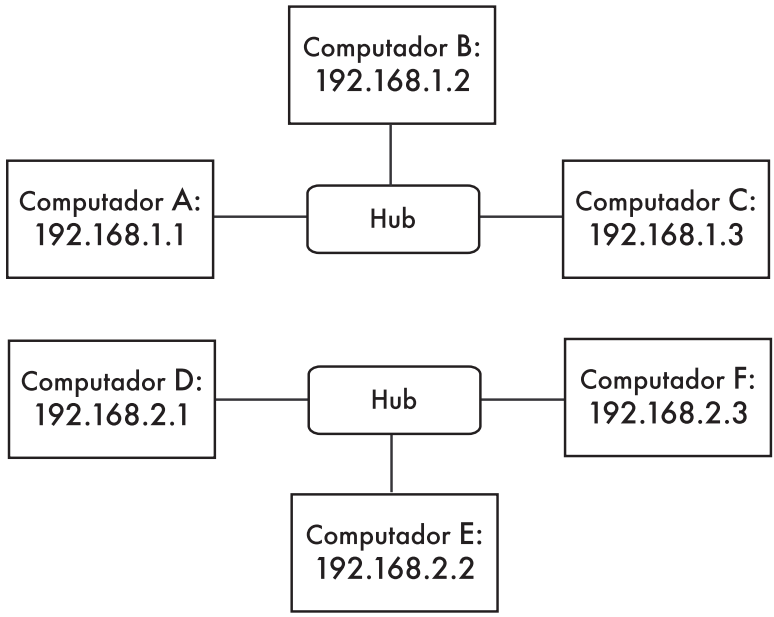


Figura 3.7: Duas redes IP separadas.

Mas como os hosts A, B e C podem comunicar-se com os hosts D, E e F? Eles necessitarão adicionar uma rota para a outra rede através do host G. Por exemplo, os hosts A-C adicionariam uma rota via 192.168.1.4. No Linux, isso é feito com o seguinte comando:

```
# ip route add 192.168.2.0/24 via 192.168.1.4
```

... e os hosts D-F adicionariam o seguinte:

```
# ip route add 192.168.1.0/24 via 192.168.2.4
```

O resultado é mostrado na **Figura 3.8**. Note que a rota é adicionada através do endereço IP no host G, que é o link local para a respectiva rede. O host A não poderia adicionar a rota via 192.168.2.4, mesmo que seja, fisicamente, a mesma máquina que 192.168.1.4 (host G), uma vez que este IP não é um link local.

Uma rota diz ao sistema operacional que a rede desejada não reside no link local imediato e deve ***encaminhar*** o tráfego através do roteador especificado. Se o host A quiser enviar um pacote ao host F, ele precisará, primeiro, enviá-lo ao host G. O host G irá, então, procurar pelo host F em sua tabela de roteamento, verificando que ele tem uma conexão direta para a rede do host F. Finalmente, o host G irá descobrir o endereço de hardware (MAC) do host F e encaminhar a ele o pacote.

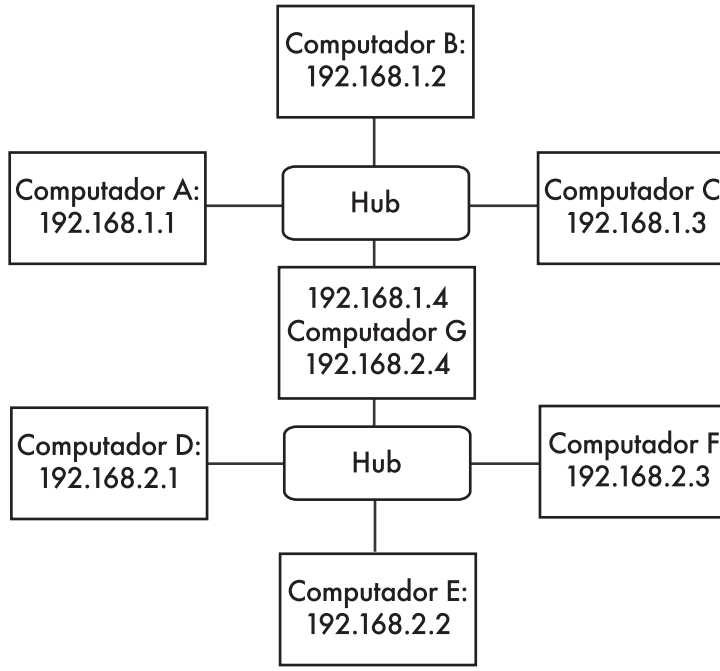
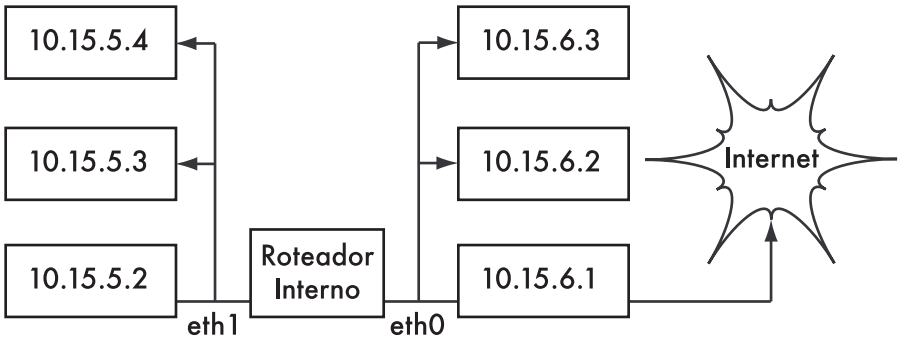


Figura 3.8: O host G atua como um roteador entre as duas redes.

Este é um exemplo muito simples de roteamento, onde o destino está apenas um intermediário (***hop***) distante da fonte. Conforme as redes tornam-se mais complexas, muitos hops precisam ser atravessados para que o destino final seja alcançado. Uma vez que não é prático, para cada máquina na Internet, conhecer a rota para cada uma das demais, fazemos uso de uma diretiva de roteamento conhecida por ***rota padrão*** (***default route***, ou ***default gateway***). Quando um roteador recebe um pacote destinado a uma rede para a qual não há uma rota específica, o pacote é encaminhado para a sua rota padrão.

A rota padrão é, tipicamente, a melhor rota de saída em sua rede, normalmente em direção a seu provedor de acesso à Internet. Um exemplo de roteador que usa uma rota padrão é mostrado na **Figura 3.9**.

Rotas podem ser atualizadas manualmente ou podem reagir automaticamente a quedas de rede e outros eventos. Alguns exemplos de protocolos de roteamento dinâmico são RIP, OSPF, BGP e OLSR. A configuração de rotas dinâmicas não faz parte do escopo deste livro mas, para leituras adicionais sobre o assunto, consulte os recursos do **Apêndice A**.




```
Tabela de roteamento para o roteador interno:

Destino         Roteador    MáscaraGen.    Opções Ref    Iface
10.15.5.0       *           255.255.255.0  U      0        eth1
10.15.6.0       *           255.255.255.0  U      0        eth0
default         10.15.6.1   0.0.0.0        UG     0        eth0
```

Figura 3.9: Quando não há rota explícita para um destino em particular, um computador deve usar a rota padrão definida em sua tabela de roteamento.

Network Address Translation (NAT)

Para que se possa acessar servidores na Internet, os endereços IP do tipo RFC1918 devem ser convertidos para endereços IP globais, publicamente roteáveis na Internet. Isto é possível através de uma técnica conhecida como ***Network Address Translation***, ou ***NAT*** (Tradução de Endereços de Rede). Um dispositivo NAT é um roteador que modifica os endereços dos pacotes ao invés de simplesmente encaminhá-los. Em um roteador NAT, a conexão com a Internet usa um (ou mais) endereço IP globalmente roteável, enquanto a rede privada usa um endereço IP do espaço de endereços privados do RFC1918. O roteador NAT permite que endereços globais possam ser compartilhados com os usuários internos que utilizam endereços privados. Ele converte os pacotes de uma forma de endereçamento para a outra quando os pacotes passam por ele. Do ponto de vista dos usuários da rede, eles estão diretamente conectados à Internet, não necessitando de nenhum software ou driver específico. Eles simplesmente usam o roteador NAT como sua rota padrão e endereçam os pacotes como fariam normalmente. O roteador NAT traduz os pacotes que estão deixando a rede privada para que usem um endereço IP global, e os traduzem novamente para um endereço interno quando retornam.

A principal conseqüência da utilização do NAT é que as máquinas que estão na Internet não têm acesso fácil aos servidores internos à organização, a não ser que sejam configuradas regras explícitas de ***encaminhamento*** (*forwarding*) no roteador. As conexões que partem de dentro do espaço de endereços privados, porém, não têm dificuldade em acessar a Internet, ainda que algumas aplicações (como voz sobre IP e alguns softwares de VPN) tenham alguns problemas ao lidar com o NAT.

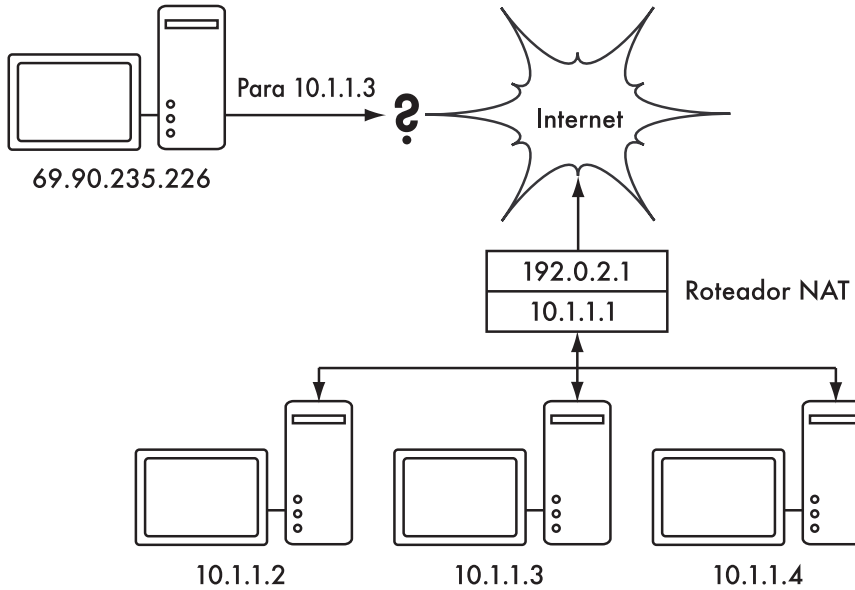


Figura 3.10: A tradução de endereços de rede (NAT) permite que você compartilhe um único endereço IP público com muitos servidores internos, mas pode dificultar o funcionamento de alguns serviços.

Dependendo de seu ponto de vista, isto pode ser considerado um defeito, um *bug* (uma vez que isto torna mais difícil a configuração de uma comunicação de mão dupla) ou um benefício, um *feature* (uma vez que, efetivamente, esta técnica implementa, sem custo adicional, um *firewall* para a sua organização inteira). Os endereços RFC1918 devem ser filtrados no limite de sua rede a fim de prevenir a entrada ou saída de tráfego RFC1918, tanto de forma acidental quanto maliciosa. Mesmo que o NAT execute algumas funções ao estilo de um firewall, ele não substitui um firewall de verdade.

A suíte de protocolos Internet

Máquinas na Internet utilizam o Internet Protocol (IP) para alcançarem umas às outras, mesmo quando estiverem separadas por máquinas intermediárias. Há mais protocolos que são usados em conjunto com o IP, provendo funções tão críticas às operações normais quanto o próprio IP. Cada pacote especifica um número de protocolo que o irá qualificar. Os protocolos mais comumente usados são o ***Transmission Control Protocol*** (***TCP***, número 6), o ***User Datagram Protocol*** (***UDP*** , número 17) e o ***Internet Control Message Protocol*** (***ICMP***, número 1). Em grupo, estes protocolos (e outros) são conhecidos como a suíte de protocolos Internet (***Internet Protocol Suite***) ou, simplesmente, ***TCP/IP***.

Os protocolos TCP e UDP introduzem o conceito de numeração de portas. Os números de portas (port numbers) permitem que vários serviços possam ser executados em um único endereço IP e, ainda assim, de forma distinta um do outro. Cada pacote tem um número de porta para a sua fonte e seu destino. Alguns números de portas são padrões bem definidos, usados para alcançar serviços bem conhecidos, como servidores de email e web. Por exemplo, um

servidor web normalmente escuta (***listen***) a porta 80, e servidores de email SMTP escutam a porta 25. Quando dizemos que um serviço "escuta" em uma porta (como a porta 80), queremos dizer que ele irá aceitar pacotes que usem seu IP como o endereço de destino, e a porta 80 como a porta de destino. Os servidores normalmente não se importam com a origem do pacote (seja seu IP ou porta), ainda que algumas vezes utilizem esta informação para estabelecer a identidade do servidor que os enviou. Ao enviar uma resposta para tais pacotes, o servidor irá usar seu próprio IP e porta como a nova origem (no caso, a porta 80).

Quando um cliente conecta-se a um serviço, ele pode utilizar como porta de origem qualquer uma, de seu lado, que não esteja sendo usada, mas deve conectar-se à porta apropriada no servidor que está provendo tal serviço (isto é, 80 para web, 25 para email). O TCP é um protocolo ***orientado à sessão*** (***session oriented***) com funções que garantem a entrega e a transmissão de pacotes (como a detecção e atenuação de congestionamento de rede, repetição de tentativas de envio, reordenação e montagem de pacotes, etc). O UDP é projetado para o fluxo de informação sem controle de conexão (***connectionless***), e não dá nenhuma garantia de entrega, ou nenhuma ordenação em particular.

O protocolo ICMP é projetado para o diagnóstico de problemas e manutenção na Internet. Ao invés de usar números de portas, ele possui tipos de mensagens (***message types***), que também são números. Diferentes tipos de mensagens são usados para solicitar uma simples resposta de um outro computador (*echo request*), notificar o remetente de uma mensagem de um possível "loop" na rota (*time exceeded*) ou informar ao remetente que um pacote não pôde ser entregue em função de regras de firewall ou outro problema (*destination unreachable*).

Até aqui você já deve ter um bom entendimento da forma como os computadores são endereçados na rede e como a informação flui entre eles. Agora, vamos dar uma rápida olhada no hardware que implementa estes protocolos de rede.

Ethernet

Ethernet é o nome do mais conhecido padrão para a conexão de computadores em uma rede local (***LAN—Local Area Network***). Ele é, algumas vezes, usado para conectar computadores individuais à Internet, através de um roteador, modem ADSL ou dispositivo wireless. Entretanto, ao conectar um único computador à Internet, pode ser que você sequer use Ethernet. O nome vem do conceito físico do éter (*ether*), o meio que foi, certa vez, considerado o responsável por carregar ondas de luz pelo espaço. O padrão oficial é chamado IEEE 802.3.

O padrão mais comum Ethernet é o 100baseT. Isto define uma taxa de transmissão de dados de 100 megabits por segundo, em um par trançado de fios, com conectores modulares RJ-45 na ponta. A topologia da rede é uma estrela, com switches ou hubs no centro da estrela e nós finais (dispositivos e switches adicionais) nas extremidades.

Endereço MAC

Cada dispositivo conectado a uma rede Ethernet tem um endereço MAC único, atribuído pelo fabricante do cartão de rede. Sua função é similar a de um endereço IP, uma vez que serve como o identificador individual que permite a um dispositivo conversar com outro. Entretanto, o escopo de um endereço MAC está limitado a um domínio de *broadcast*, definido como todos os computadores conectados fisicamente por cabos, hubs, switches e bridges, sem cruzar roteadores ou gateways de Internet. Os endereços MAC nunca são usados diretamente na Internet e não são transmitidos além dos roteadores.

Hubs

Os ***hubs*** Ethernet conectam múltiplos dispositivos Ethernet de par trançado entre si. Eles trabalham na camada física (a camada mais baixa, primeira). Eles repetem os sinais recebidos em cada porta[2] para todas as demais. Assim, os hubs podem ser considerados como simples repetidores. Em função deste projeto, apenas uma porta pode transmitir por vez. Caso dois dispositivos transmitam ao mesmo tempo, eles corrompem a transmissão um do outro, devendo ambos cancelar sua transmissão e retransmitir, posteriormente, os pacotes. Isto é conhecido como uma ***colisão*** (***collision***), e cada servidor fica responsável por detectar as colisões durante uma transmissão e pela retransmissão de seus próprios pacotes, quando necessário.

Quando um excesso de colisões é detectado em uma porta, alguns hubs podem desconectar (***partition***) tal porta por algum tempo, limitando o impacto do problema no resto da rede. Quando uma porta é desconectada, os dispositivos ligados a ela não mais podem comunicar-se com o restante da rede. Redes baseadas em hubs são, geralmente, mais robustas que as que utilizam conexões Ethernet coaxiais (também conhecidas como 10base2 ou ThinNet), onde dispositivos com problemas podem indisponibilizar todo um segmento de rede. Mas os hubs têm limitações em sua utilidade, uma vez que podem tornar-se, facilmente, pontos de congestionamento em redes de alto tráfego.

Switches

Um ***switch*** é um dispositivo que opera de forma parecida a um hub, mas que fornece uma conexão dedicada (chaveada, ***switched***) entre as suas portas. Ao invés de repetir todo o tráfego em todas as portas, o switch determina quais portas estão se comunicando diretamente e, temporariamente, as conecta. Em geral, os switches oferecem um desempenho muito melhor que os hubs, especialmente em redes de alto tráfego, com muitos computadores. Eles não são muito mais caros que os hubs e os estão substituindo em muitas situações.

Os switches trabalham na camada de comunicação de dados (a segunda camada), uma vez que eles interpretam e atuam no endereço MAC dos pacotes que recebem. Quando um pacote chega à porta de um switch, o mesmo anota a fonte do endereço MAC, que fica associado àquela porta. A informação é

---

2. N. do T. - Quando falamos em portas aqui, o leitor deve observar que elas não são do mesmo tipo daquelas a que nos referimos no protocolo TCP/IP. Aqui estamos falando de portas que correspondem ao conector físico de um hub. O mesmo ocorrerá quando falarmos em portas de um switch. Normalmente será fácil de distinguir, neste texto, a qual tipo de porta estamos nos referindo.

armazenada em uma ***MAC table*** (tabela de endereços MAC), internamente. Para cada pacote que recebe, o switch verifica qual o endereço MAC destino em sua MAC table e transmite o pacote para a porta correspondente. Caso o endereço MAC não seja encontrado na MAC table, o pacote é transmitido para todas as interfaces. Caso o endereço de destino corresponda à mesma porta pela qual ele foi enviado, o pacote é filtrado e não é encaminhado para essa porta.

Hubs versus Switches

Hubs são considerados dispositivos pouco sofisticados, uma vez que eles retransmitem, de forma ineficiente, todo o tráfego em todas as portas. Esta simplicidade leva tanto a um desempenho fraco quanto a um problema de segurança. O desempenho é fraco porque a largura de banda deve ser dividida entre todas as suas portas. Uma vez que o tráfego é visto por todas as portas, qualquer servidor na rede pode monitorá-lo integralmente.

Os switches criam conexões virtuais entre as portas que estão transmitindo e recebendo. Isto leva a um melhor desempenho porque muitas conexões virtuais podem ser estabelecidas simultaneamente. Switches mais sofisticados (e caros) podem direcionar melhor o tráfego através da inspeção dos pacotes em níveis mais altos (como as camadas de transporte e aplicação), permitindo a criação de VLANs e implementando outras funcionalidades avançadas.

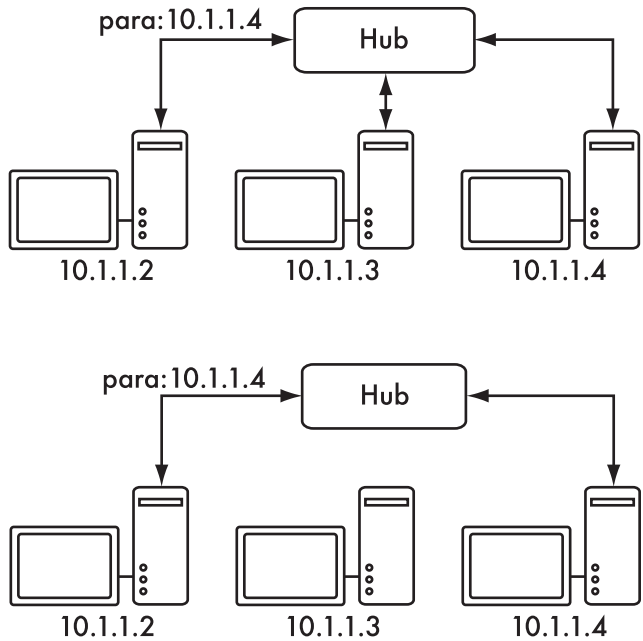


Figura 3.11: Um hub simplesmente repete todo o tráfego para todas as portas, enquanto um switch estabelece uma conexão temporária dedicada entre as portas que necessitam comunicar-se.

Um hub pode ser usado onde a repetição do tráfego em todas as portas for desejável. Por exemplo, quando, explicitamente, você permite que uma máquina

de monitoramento inspecione o tráfego de toda a rede. Muitos switches fornecem uma porta de monitoração (***monitor port***) que permite a repetição do tráfego nela, especificamente para esta função.

Os hubs são mais baratos que os switches. Entretanto, o preço dos switches caiu dramaticamente com o passar do tempo. Desta forma, os hubs de redes antigas devem ser substituídos, sempre que possível, por novos switches.

Tanto hubs como switches podem oferecer serviços gerenciados (***managed***). Alguns destes serviços incluem a habilidade de configurar a velocidade do link (10baseT, 100baseT, 1000baseT, *full* ou *half duplex*) por porta, habilitar gatilhos (*triggers*) para inspecionar eventos de rede (como mudanças de endereço MAC ou pacotes com má formação) e, usualmente, incluem bilhetagem de portas (*port counters*) para facilitar as estatísticas de utilização de banda. Um switch gerenciado que forneça informações sobre a quantidade de dados que entram ou saem de cada porta física pode simplificar bastante a gestão da rede. Estes serviços são, tipicamente, disponibilizados via SNMP, ou podem ser acessados via telnet, ssh, inteface web ou alguma ferramenta customizada de configuração.

Roteadores e firewalls

Enquanto hubs e switches fornecem conectividade para um segmento local de rede, a função de um roteador é a de encaminhar pacotes entre diferentes segmentos de rede. Um roteador, tipicamente, tem duas ou mais interfaces físicas de rede. Ele pode incluir o suporte a diferentes tipos de rede, como Ethernet, ATM, DSL ou conexão discada. Os roteadores podem ser dispositivos de hardware dedicados (como os roteadores Cisco ou Juniper) ou podem ser feitos a partir de um PC padrão, com múltiplos cartões de rede e o software apropriado.

Roteadores localizam-se no limite (***edge***) entre duas ou mais redes. Por definição, eles têm uma conexão para cada rede e, como são máquinas de fronteira, podem ter outras responsabilidades além do roteamento. Muitos roteadores executam funções de ***firewall***, provendo um mecanismo para a filtragem ou redirecionamento de pacotes que não se enquadram na política de acesso ou segurança da rede. Eles também podem fornecer serviços de tradução de endereços (NAT).

Os roteadores possuem uma variação muito grande de custo e funcionalidades. Os mais baratos e menos flexíveis são dispositivos de hardware simples e dedicados, freqüentemente com funcionalidade NAT, utilizados para compartilhar uma conexão Internet com alguns poucos computadores. O próximo passo é um roteador por software, que consiste em um sistema operacional rodando em um PC com múltiplas interfaces de rede. Sistemas operacionais padrão como o Microsoft Windows, Linux ou BSD realizam funções de roteamento e são muito mais flexíveis que os dispositivos de hardware de baixo custo. Entretanto, eles têm os mesmos problemas que os PCs convencionais: alto consumo de energia, grande e complexo número de componentes não confiáveis e maior necessidade de configuração.

Os mais caros são roteadores de hardware de alto nível, feitos por empresas como a Cisco ou a Juniper. Eles tendem a ter um desempenho muito superior, mais funcionalidade e confiabilidade muito maior que roteadores por

software implementados com PCs. Também é possível adquirir suporte técnico e ter contratos de manutenção para eles.

A maioria dos roteadores modernos oferecem mecanismos para monitorar e registrar seu desempenho remotamente, normalmente através do ***Simple Network Management Protocol*** (***SNMP***—Protocolo Simples de Gerenciamento de Rede), ainda que os dispositivos mais baratos freqüentemente omitam esta funcionalidade.

Outros equipamentos

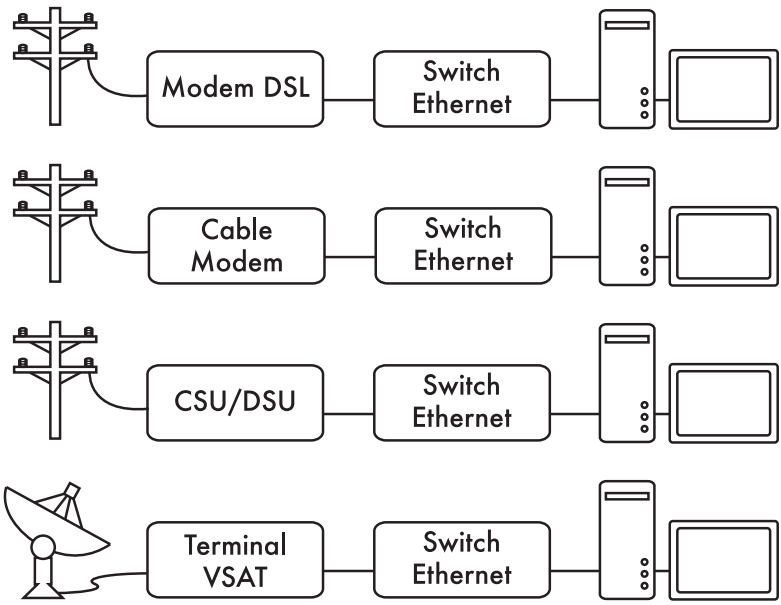


Figura 3.12: Muitos modems DSL, cable modems, CSU/DSUs, pontos de acesso wireless e terminais VSAT fornecem conexão para um cabo Ethernet.

Cada rede física possui um equipamento terminal associado. Por exemplo, conexões VSAT são compostas de um prato de satélite conectado a um terminal que, ou liga-se a uma placa dentro de um PC, ou fornece uma conexão Ethernet padrão. Linhas DSL usam um ***modem DSL*** que conecta uma linha telefônica a um dispositivo local, seja uma rede Ethernet ou um único computador via USB. ***Cable modems*** ligam um cabo de televisão à Ethernet, ou a algum tipo de cartão no PC. Alguns tipos de circuitos de telecomunicação (como T1 ou T3) usam uma CSU/DSU para a conexão com uma porta serial ou Ethernet. Linhas da rede de telefonia pública usam modems para conectar um computador ao telefone, normalmente através de um cartão interno ou uma porta serial. Existem ainda muitos tipos diferentes de equipamentos de rede sem fio que conectam-se a uma variedade de rádios e antenas, mas praticamente todos eles fornecem uma conexão para um cabo Ethernet.

Conectando todas as coisas

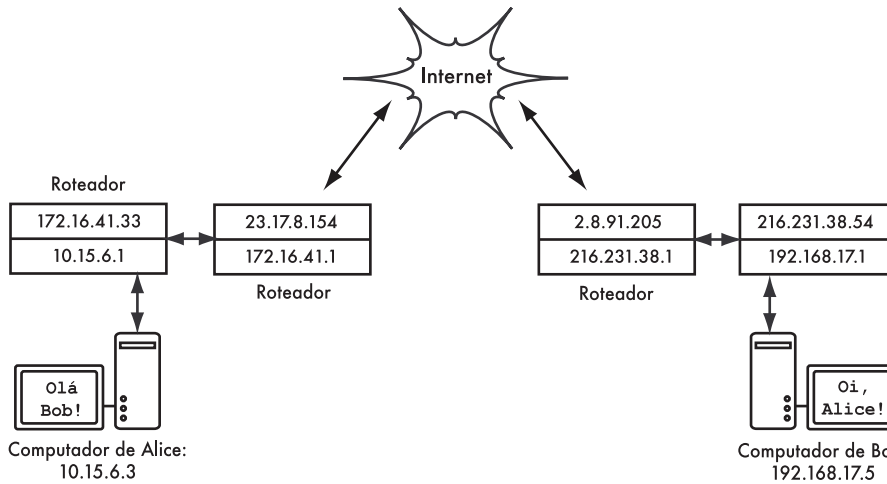


Figure 3.13: Rede Internet: Cada segmento de rede tem seu próprio roteador com dois endereços IP, fazendo seu "link local" para duas redes diferentes. Os pacotes são encaminhados entre os roteadores até que atinjam seu destino final.

Uma vez que todos os nós de rede tenham endereços IP, eles podem enviar e receber pacotes pelos endereços uns dos outros. Através de roteamento e encaminhamento, estes pacotes podem atingir redes que não estejam fisicamente conectadas ao nó que os originou. Este processo descreve muito do que acontece na Internet.

Neste exemplo, você pode ver o caminho que os pacotes percorrem quando Alice conversa com Bob através de um serviço de mensagens instantâneas. Cada linha representa um cabo Ethernet, uma conexão wireless, ou algum outro tipo de rede física. O símbolo da nuvem é comumente usado para representar "A Internet", significando qualquer rede IP que está no caminho da comunicação. Alice ou Bob não precisam preocupar-se com a forma como a rede funciona, desde que os roteadores encaminhem o tráfego a seus destinos. Não fosse por causa dos protocolos de Internet e da colaboração entre todos os elementos da rede, este tipo de comunicação seria impossível.

Projetando a rede física

Pode parecer estranho falar de uma rede "física" quando estamos construindo redes sem fio. Afinal, qual é a parte física de uma rede? Em redes wireless, o meio físico que usamos para a comunicação é, obviamente, a energia eletromagnética. Mas, no contexto deste capítulo, a rede física refere-se ao tópico mundano de "onde colocamos as coisas". Como organizamos o equipamento de maneira que possamos alcançar nossos clientes wireless? Seja em um prédio de escritórios ou espalhadas por muitos quilômetros, redes sem

fio estão naturalmente implementadas nestas três configurações lógicas: ***links ponto-a-ponto***, ***links ponto-para-multiponto*** e ***nuvens multiponto-para-multiponto***. Enquanto porções diferentes de sua rede possam tomar vantagem de todas estas três configurações, qualquer link individual estará em uma destas três topologias.

Ponto-a-ponto

Links ***ponto-a-ponto*** tipicamente fornecem uma conexão à Internet onde é impossível o acesso de outra forma. Um lado da conexão ponto-a-ponto já tem uma conexão com a Internet, enquanto o outro usará este link para alcançá-la. Por exemplo, uma Universidade pode ter uma conexão de alta velocidade com a Internet, utilizando *frame relay* ou VSAT no campus central, mas o custo seria alto demais para ter o mesmo tipo de conexão em um prédio localizado fora de seus limites. Caso o prédio principal tenha uma visão desobstruída para a localidade remota, uma conexão ponto-a-ponto pode ser utilizada. Isto pode aumentar, ou mesmo substituir, conexões discadas existentes. Com antenas apropriadas e uma linha de visão clara, links ponto-a-ponto podem ser estabelecidos para distâncias superiores a 30 quilômetros.

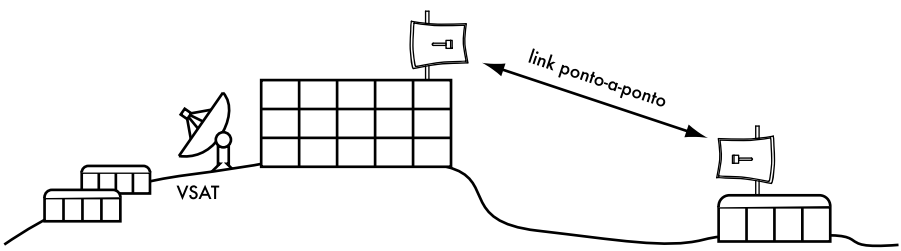


Figura 3.14: Um link ponto-a-ponto permite a uma localidade remota compartilhar uma conexão à Internet no prédio principal.

É claro que, uma vez que uma conexão ponto-a-ponto é estabelecida, outras podem ser usadas para estender a rede para distâncias maiores. Caso o prédio remoto de nosso exemplo esteja no topo de uma colina, ele pode ter visibilidade para outras localizações importantes que não poderiam ser acessadas diretamente do campus central. Com a instalação de outro link ponto-a-ponto na localidade remota, mais um nó poderia juntar-se à rede, usando a mesma conexão à Internet do campus central.

Links ponto-a-ponto não necessariamente precisam envolver o acesso à Internet. Imagine que você tenha que, fisicamente, dirigir até uma estação de monitoramento meteorológico no topo de uma montanha para coletar os dados registrados nela ao longo do tempo. Você pode conectar esta estação com um link ponto-a-ponto, permitindo que a coleta e monitoração de dados ocorram em tempo real, sem que seja necessário o deslocamento até ela. Redes wireless fornecem largura de banda suficiente para o transporte de grande quantidade de dados (incluindo áudio e vídeo) entre quaisquer dois pontos conectados entre si, mesmo que não exista uma conexão direta com a Internet.

Ponto-para-multiponto

O segundo tipo de rede mais encontrado é o ***ponto-para-multiponto***. Sempre que vários nós[3] estão em comunicação com um ponto de acesso central temos uma aplicação de ponto-para-multiponto. O típico exemplo de um leiaute ponto-para-multiponto é o uso de um access point que provê a conexão para vários laptops. Os laptops não se comunicam entre si diretamente, mas devem estar nas proximidades do access point para que possam utilizar a rede.

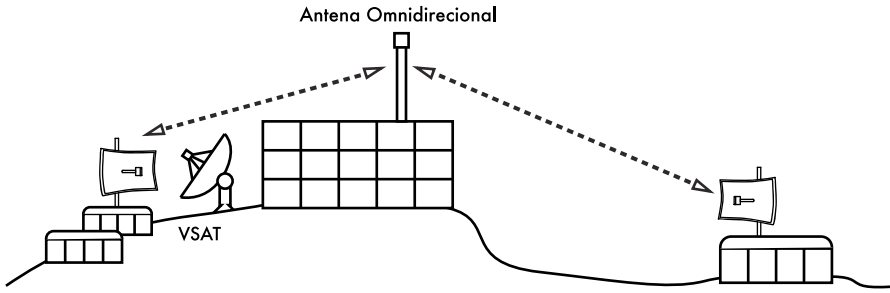


Figura 3.15: A conexão VSAT central é, agora, compartilhada por múltiplas localidades remotas. Todas as três localidades podem também comunicar-se entre si a velocidades muito maiores que o VSAT.

A rede ponto-para-multiponto pode também ser aplicada ao nosso exemplo anterior, na Universidade. Suponha que o edifício remoto, ao topo de uma colina, está conectado ao campus central por um link ponto-a-ponto. Ao invés de configurar vários links ponto-a-ponto para distribuir a conexão com a Internet, uma simples antena, visível a partir dos vários prédios, pode ser utilizada. Este é um clássico exemplo de uma conexão de um ***ponto*** de rede ampla (a localidade remota no topo da montanha) ***para multiponto*** (muitas localidades no vale sob a montanha).

Note que há uma série de considerações de desempenho que devem ser consideradas quando se utiliza a conexão ponto-para-multiponto com distâncias muito longas, que serão tratadas posteriormente neste capítulo. Tais links são possíveis e úteis em muitas circunstâncias, mas não cometa o clássico erro de instalar uma única torre de rádio no meio da cidade esperando atender a milhares de clientes, como você poderia fazer com uma estação de rádio FM. Como veremos adiante, redes de dados bidirecionais comportam-se de maneira bem diferente da transmissão *broadcast* de uma rádio.

Multiponto-para-multiponto

O terceiro tipo de leiaute de rede é o ***multiponto-para-multiponto***, também chamado de rede ***ad-hoc*** ou ***mesh***. Em uma rede multiponto-para-multiponto, não existe uma autoridade central. Todos os nós da rede encarregam-se do tráfego um do outro, conforme o necessário, e cada nó comunica-se com o outro diretamente.

---

3. Um **nó** é qualquer dispositivo capaz de enviar e receber dados em uma rede. Access points, roteadores, computadores e laptops são exemplos de nós.

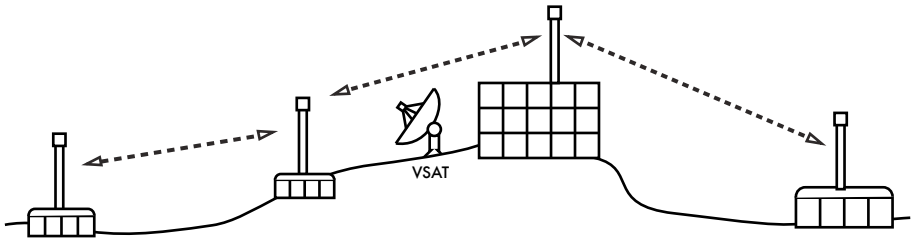


Figura 3.16: Uma rede mesh, multiponto-para-multiponto. Cada ponto pode atingir o outro a uma velocidade altíssima, ou usar o VSAT central para a conexão com a Internet.

O benefício do leiaute desta rede é que, mesmo que nenhum dos nós esteja ao alcance de um ponto de acesso central, ainda assim eles podem comunicar-se um com o outro. Boas implementações de redes mesh são "auto-curáveis", o que significa que elas detectam automaticamente problemas de roteamento e os consertam quando necessário. A extensão de uma rede mesh é simplesmente feita com a adição de mais nós. Caso um dos nós na "nuvem" seja um gateway para a Internet, esta conexão pode ser compartilhada por todos os clientes da rede.

Duas grandes desvantagens desta topologia são a complexidade aumentada e o desempenho diminuído. A segurança neste tipo de rede também é uma preocupação, uma vez que todos os participantes carregam o tráfego, um do outro. Redes multiponto-para-multiponto tendem a ser difíceis de diagnosticar devido ao grande número de variáveis, como os nós que entram e deixam a rede. Nuvens multiponto-para-multiponto têm, tipicamente, uma capacidade reduzida se comparadas com redes ponto-a-ponto ou ponto-para-multiponto, em função da sobrecarga adicional da gestão do roteamento de rede e contenção no espectro de rádio.

De qualquer forma, as redes mesh são úteis em muitas circunstâncias. Mais adiante, nesse capítulo, veremos um exemplo de construção de uma rede mesh multiponto-para-multiponto utilizando um protocolo de roteamento chamado OLSR.

Use a tecnologia adequada

Todos estes esquemas de rede podem ser usados de forma complementar em uma grande rede e, obviamente, pode-se usar também técnicas tradicionais de redes cabeadas sempre que possível. É prática comum, por exemplo, usar um link wireless de longa distância para prover acesso à Internet para uma localidade remota e, a partir daí, distribuir pontos de acesso sem fio locais para distribuir a conexão. Um dos clientes deste ponto de acesso pode também atuar como um nó mesh, permitindo que a rede espalhe-se organicamente entre usuários de laptops. Todos, em última instância, estão usando o link ponto-a-ponto original para o acesso à Internet.

Agora que temos uma clara idéia de como as redes wireless estão tipicamente organizadas, podemos começar a entender como a comunicação é possível nestas redes.

Redes wireless 802.11

Antes que os pacotes possam ser encaminhados e roteados para a Internet, as camadas um (física) e dois (o link de dados) precisam estar conectadas. Sem conexão ao link local, os nós da rede não podem comunicar-se entre si e rotear pacotes.

Para prover conectividade física, os dispositivos de rede wireless devem operar na mesma porção do espectro de rádio. Como vimos no **Capítulo 2**, isto significa que rádios 802.11a irão se comunicar com rádios 802.11a numa freqüência próxima a 5 GHz, e rádios 802.11b/g irão se comunicar com outros rádios 802.11b/g na faixa de 2.4 GHz. Mas um dispositivo 802.11a não irá interoperar com um dispositivo 802.11b/g, uma vez que eles utilizam porções completamente diferentes do espectro eletromagnético.

Mais especificamente, cartões wireless devem estar de acordo sobre o canal comum que utilizarão. Se um cartão de rádio 802.11b está configurado para usar o canal 2, enquanto outro está configurado para o canal 11, eles não falarão entre si.

Quando dois cartões wireless estão configurados para usar o mesmo protocolo, no mesmo canal de rádio, então eles estão prontos para negociar a conectividade da camada de comunicação de dados. Cada dispositivo 802.11a/b/g pode operar em um destes quatro possíveis modos:

1. ***Modo master*** (também chamado de ***AP*** ou ***modo de infra-estrutura***) é usado para criar um serviço que se parece com um ponto de acesso tradicional. O cartão wireless cria uma rede com um nome específico (chamado SSID) e canal, oferecendo serviços de rede nele. No modo master, os cartões wireless gerenciam toda a comunicação relativa à rede (autenticando clientes wireless, tratando da contenção do canal, repetindo pacotes, etc). Cartões wireless em modo master podem apenas comunicar-se com cartões associados a ele em modo gerenciado.

2. ***Modo gerenciado*** é chamado também, algumas vezes, de modo ***cliente***. Cartões wireless no modo gerenciado irão unir-se a uma rede criada pelo master, automaticamente trocando seu canal para corresponder a ele. Eles então apresentam qualquer credencial que é necessária para o master e, se estas credenciais são aceitas, diz-se que eles estão ***associados*** ao master. Cartões no modo gerenciado não se comunicam diretamente um com o outro e se comunicarão apenas com o master associado.

3. ***Modo ad-hoc*** cria uma rede multiponto-para-multiponto, onde não existe um único nó master ou AP. Em modo ad-hoc, cada cartão wireless comunica-se diretamente com os vizinhos. Os nós devem estar ao alcance para que se comuniquem e devem estar de acordo quanto ao nome da rede e o canal utilizado.

4. ***Modo monitor*** é usado por algumas ferramentas (tais como **Kismet**, veja no **Capítulo 6**) para passivamente inspecionar todo o tráfego de rádio em um dado canal. Quando estão no modo monitor, os cartões

wireless não transmitem nenhum dado. Isto é útil para a análise de problemas em um link wireless ou para observar a utilização do espectro na área monitorada. O modo monitor não é usado para a comunicação normal.

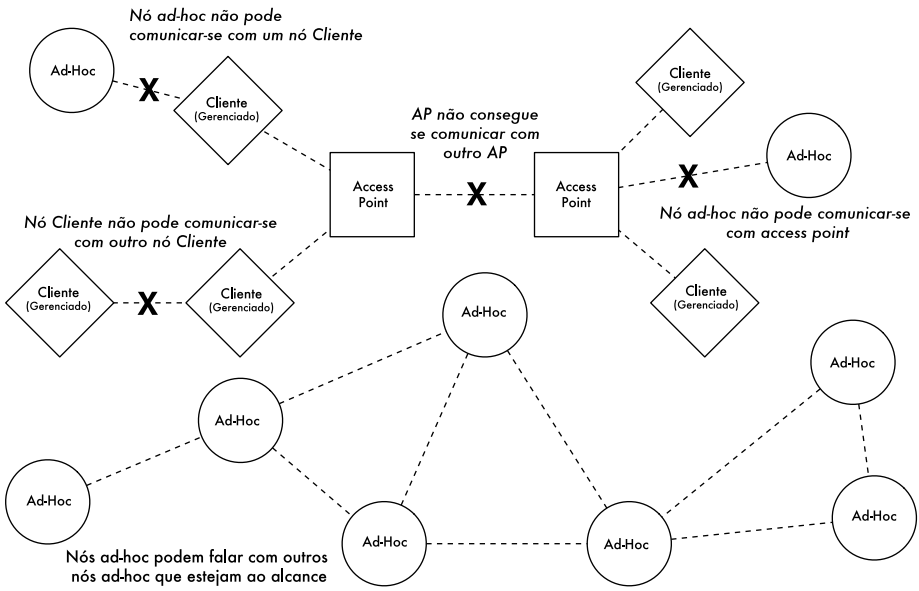


Figura 3.17: Nós APs, Clientes e Ad-Hoc

Quando se implementa um link ponto-a-ponto ou ponto-para-multiponto, um rádio irá operar, tipicamente, em modo master, enquanto os demais irão operar em modo gerenciado. Em um mesh multiponto-para-multiponto, todos os rádios operam em modo ad-hoc e, assim, podem comunicar-se diretamente uns com os outros.

É importante manter estes modos em mente quando estiver projetando sua rede. Lembre-se que clientes em modo gerenciado não podem comunicar-se entre si diretamente, assim, é provável que você queira implementar um site de repetidores no modo master ou ad-hoc. Como veremos adiante neste capítulo, o modo ad-hoc é mais flexível, mas tem uma série de problemas de desempenho quando comparado ao uso dos modos master ou gerenciado.

Rede Mesh com OLSR

A maioria das redes Wi-Fi opera em modo de infra-estrutura, ou seja, consiste de um access point em algum lugar (com um rádio operando em modo master), conectado a uma linha DSL ou algum outro tipo de rede cabeada de larga escala. Em um ***hotspot*** deste tipo, o access point usualmente atua como uma estação master, distribuindo o acesso Internet a seus clientes, que operam em modo gerenciado. Esta topologia é similar a um serviço celular GSM. Os telefones celulares conectam-se a uma estação base—sem a presença desta

estação, os celulares não podem comunicar-se entre si. Se você quiser passar um trote para o seu amigo sentado à sua frente em uma mesa, seu celular envia os dados para uma estação base de sua operadora, que pode estar a três quilômetros de distância, e esta, então, envia os dados para o telefone de seu amigo.

Da mesma forma, cartões Wi-Fi em modo gerenciado não podem comunicar-se diretamente. Clientes—por exemplo, dois laptops em uma mesma mesa—precisam usar um access point como um ponto de passagem (relay). Qualquer tráfego entre os clientes conectados a um access point tem que ser enviado duas vezes. Se o cliente A e o C comunicam-se, o cliente A envia os dados para o access point B, e então o access point retransmitirá os dados para o cliente C. Uma simples transmissão pode ter a velocidade de 600 kByte/s (isto é aproximadamente a máxima velocidade que você alcança com um 802.11b) em nosso exemplo—então, como os dados tem que ser repetidos pelo access point para alcançar seu alvo, a velocidade efetiva entre dois clientes será de apenas 300 kByte/s.

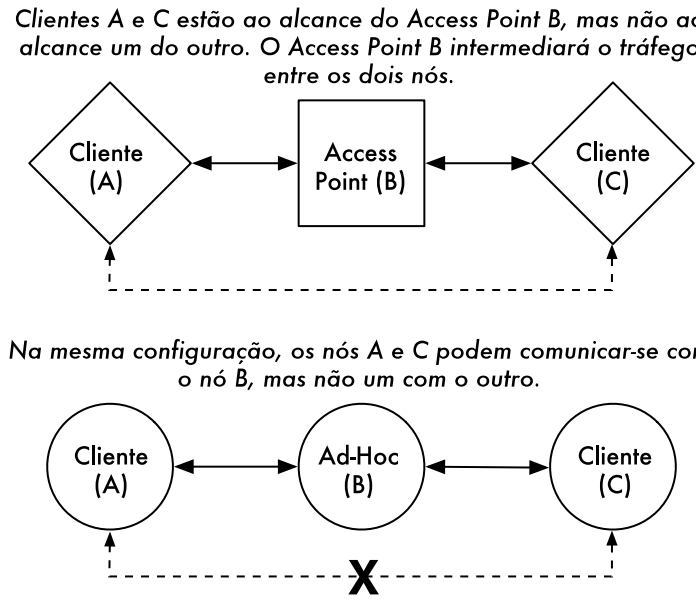


Figura 3.18: O Access Point B irá intermediar o tráfego entre os clientes A e C. Em modo Ad-Hoc, o nó B, por padrão, não irá intermediar o tráfego entre A e C.

No modo ad-hoc não existe uma relação hierárquica master-cliente. Os nós podem comunicar-se diretamente, desde que estejam ao alcance de suas interfaces wireless. Assim, em nosso exemplo, ambos os computadores poderiam atingir a velocidade plena de transmissão de dados trabalhando em modo ad-hoc, em circunstâncias ideais.

A desvantagem do modo ad-hoc é que os clientes não repetem o tráfego destinado a outros clientes. No exemplo do access point, os dois clientes A e C podem não estar ao alcance um do outro, mas podem comunicar-se desde que ambos estejam ao alcance do access point.

Como padrão, nós ad-hoc não repetem dados, mas eles podem fazê-lo se aplicarmos ***roteamento***. As redes mesh são baseadas na estratégia de que cada nó em modo mesh atua como um repetidor para estender a cobertura da rede wireless. Quanto mais nós, melhor a cobertura de rádio e maior o alcance da nuvem mesh.

Há um considerável ponto negativo que mencionaremos agora. Caso o dispositivo utilize apenas uma interface de rádio, a capacidade de banda é significativamente reduzida a cada vez que o tráfego é repetido por nós intermediários no caminho entre A e B. Haverá também interferências na transmissão, uma vez que os nós compartilham o mesmo canal. Assim, redes ad-hoc de baixo custo podem prover boa cobertura de rádio nos últimos quilômetros de uma rede wireless comunitária, ao custo de velocidade—especialmente se a densidade dos nós e a potência de transmissão for alta.

Se uma rede ad-hoc consiste de apenas alguns poucos nós que estão constantemente ligados, não se deslocam e tem sempre links estáveis de rádio —ou seja, uma longa lista de "se"—então é possível configurar tabelas de roteamento para cada nó, manualmente.

Infelizmente, essas condições raramente existem no mundo real. Nós podem falhar, dispositivos com Wi-Fi circulam por todos os lados e a interferência pode derrubar links de rádio a qualquer momento. E ninguém quer ficar atualizando manualmente tabelas de roteamento a cada vez que um novo nó integra a rede. Através do uso de protocolos de roteamento que, automaticamente, mantém tabelas de roteamento individuais em cada nó envolvido, podemos evitar estes problemas. Protocolos populares de roteamento do mundo cabeado (como o OSPF) não funcionam bem neste ambiente porque não são projetados para lidar com conexões intermitentes ou topologias que mudam rapidamente.

Roteamento mesh com olsrd

O ***Optimized Link State Routing Daemon*** (***olsrd***—*http://www.olsr.org*)—traduz-se por serviço de roteamento otimizado para estado de conexão—é uma aplicação de roteamento desenvolvida para funcionar em redes wireless. Nos concentraremos neste software de roteamento por várias razões. Ele é um projeto de código aberto com suporte a Mac OS X, Windows 98, 20000, XP, Vista, Linux, FreeBSD, OpenBSD e NetBSD. Olsrd está disponível para access points que usam (ou podem usar) o Linux, como a família Linksys WRT54G, Asus WI500g, AccessCube ou Pocket PCs rodando o Familiar Linux e ele também é o padrão para os kits Metrix rodando o Pyramid. O olsrd pode trabalhar com múltiplas interfaces e ser estendido através de plugins. Suporta o protocolo IPv6 e é ativamente desenvolvido e usado em redes comunitárias em todo o mundo.

Note que há muitas implementações do *Optimized Link State Routing*, que começou como uma proposta para o IETF escrita no INRIA da França. A implementação do olsrd começou como uma tese de mestrado de Andreas Toennesen na UniK University. Com base na experiência prática de redes comunitárias livres, o serviço de roteamento foi modificado. O olsrd difere, hoje, significativamente de seu projeto original, pois passou a incluir um mecanismo chamado *Link Quality Extension* (extensão de qualidade de linha) que mede a perda de pacotes entre os nós e calcula as rotas levando em conta esta

informação. Esta extensão quebra a compatibilidade com os serviços de roteamento que seguem a especificação original do INRIA. O olsrd disponível em *www.olsr.org* pode ser configurado para que se comporte de acordo com a especificação do IETF que não possui esta funcionalidade—mas não há razão para desabilitar a *Link Quality Extension*, a não ser que a compatibilidade com outras implementações seja necessária.

Teoria

Depois que o olsrd está em execução por algum tempo, um nó sabe da existência de todos os outros nós da nuvem mesh e quais podem ser usados para rotear tráfego para eles. Cada nó mantém uma tabela de roteamento cobrindo toda a rede mesh. Este tratamento dado ao roteamento mesh é chamado de ***roteamento proativo*** (***proactive routing***). Por outro lado, algoritmos de ***roteamento reativo*** (***reactive routing***) procuram rotas apenas quando é necessário o envio de dados para um nó específico.

Há prós e contras para o roteamento proativo e existem muitas outras idéias sobre a forma de se implementar roteamento mesh que valeriam a pena mencionar. A maior vantagem do roteamento proativo é que você sabe quais são os nós que compõem a rede, não precisando esperar que rotas sejam encontradas. Uma maior sobrecarga do protocolo e maior utilização de processamento são algumas das desvantagens. Em Berlim, a comunidade Freifunk opera uma nuvem mesh onde o olsrd tem que gerenciar mais de 100 interfaces. A média de carga de CPU, causada pelo olsrd em um Linksys WRT54G, rodando a 200 MHz, é de 30% na rede mesh de Berlim. Há claramente um limite na capacidade de escala de um protocolo proativo—dependendo de quantas interfaces são utilizadas e da freqüência de atualização das tabelas de roteamento. A manutenção de rotas em uma nuvem mesh estática dá menos trabalho do que em uma onde os nós mudam constantemente de lugar, uma vez que, no primeiro caso, as tabelas de roteamento necessitam ser atualizadas com menor freqüência.

Mecanismo

Um nó rodando olsrd está, constantemente, enviando mensagem de broadcast 'Hello' (Olá) em intervalos de tempo determinados, de forma que os vizinhos possam detectar sua presença. Cada nó faz a estatística de quantos 'Hellos' foram perdidos ou recebidos de cada vizinho—obtendo, desta maneira, informações sobre a topologia e qualidade do link para os nós da vizinhança. A informação obtida sobre a topologia é transmitida como mensagens de controle de topologia (Topology Control, ou TC messages) e encaminhada pelos vizinhos que o olsrd escolheu para serem retransmissores multiponto.

O conceito de retransmissores multiponto (multipoint relays) é uma idéia nova em roteamento proativo que surgiu com o projeto do OLSR. Se cada nó retransmite a informação de topologia que recebeu, uma sobrecarga desnecessária é gerada. Tais transmissões são redundantes se um nó tem muitos vizinhos. Assim, um nó olsrd decide quais vizinhos, que são retransmissores multiponto favoráveis, irão encaminhar as mensagens de controle de topologia. Note que os retransmissores multiponto são escolhidos

para o propósito de encaminhamento de mensagens TC. A carga de trabalho é roteada considerando todos os nós disponíveis.

Há outros dois tipos de mensagens no OLSR que anunciam informação: quando um nó oferece um gateway para outras redes (mensagens HNA) ou quando o mesmo possui múltiplas interfaces (mensagens MID). Não há muito a dizer sobre estas mensagens, a não ser que elas existem. Mensagens HNA tornam o olsrd bastante conveniente quando é feita a conexão com a Internet através de um dispositivo móvel. Quando um nó mesh é movido de um lado a outro, ele irá detectar gateways para outras redes, escolhendo aquele para o qual existe a melhor rota. Entretanto, o olsrd não é infalível. Se um nó anuncia que é um gateway para a Internet—o que ele não é, seja porque nunca foi ou porque está desconectado no momento—os demais nós acreditarão nesta informação. Este pseudo-gateway é um buraco negro. Para contornar este problema, um plugin para gateway dinâmico foi escrito. Este plugin irá automaticamente detectar, no gateway, se ele está realmente conectado e se o link ainda está ativo. Caso contrário, o olsrd suspende o envio de falsas mensagens HNA. É altamente recomendável instalar e utilizar este plugin, ao invés de habilitar estaticamente mensagens HNA.

Prática

O olsrd implementa roteamento baseado em IP em uma aplicação no espaço do usuário—a instalação é relativamente fácil. Pacotes de instalação estão disponíveis para OpenWRT, AccessCube, Mac OS X, Debian GNU/Linux e Windows. O OLSR é componente padrão do Metrix Pyramid. Caso você tenha que compilar a partir do código-fonte, primeiro leia a documentação que está presente na distribuição do programa. Caso tudo esteja configurado corretamente, o que você tem a fazer é executar o olsr.

Antes de mais nada, certifique-se de que cada nó tem um endereço IP único, designado de forma estática, para cada interface utilizada na rede mesh. Não é recomendado (e também não é prático) usar DHCP em uma rede mesh baseada em IP. Um pedido de DHCP não será respondido por um servidor DHCP se o solicitante precisar passar por vários hops para chegar até ele, e aplicar a retransmissão de DHCP através de uma rede mesh é virtualmente impraticável. Esta questão poderia ser resolvida com o uso de IPv6, uma vez que ele permite espaço suficiente para a geração de endreços IP únicos a partir do endereço MAC de cada interface envolvida (como sugerido em "IPv6 Stateless Address Autoconfiguration in large mobile ad hoc networks" de K. Weniger e M. Zitterbart, 2002).

Uma página wiki, onde cada pessoa interessada possa escolher um endereço IPv4 para cada interface onde o serviço olsr esteja rodando, pode servir muito bem a esse propósito. Não há maneira fácil de automatizar o processo se o IPv4 for utilizado.

O endereço de broadcast deve ser 255.255.255.255 para as interfaces mesh em geral, como uma convenção. Não há razão para configurar o endereço de broadcast explicitamente, uma vez que o olsrd pode ser configurado para sobrescrever o endereço de broadcast por este convencionado. Deve-se certificar, apenas, que as configurações são as mesmas em todos os lugares. O oslrd pode encarregar-se disto. Quando um arquivo de configuração oslrd é

preparado e distribuído, esta funcionalidade deve estar habilitada para evitar confusões do tipo "por que os outros nós não conseguem ver a minha máquina?!?".

Agora vamos configurar a interface wireless. Aqui está um exemplo de comando de configuração de um cartão Wi-Fi com o nome wlan0 usando o Linux:

```
iwconfig wlan0 essid olsr.org mode ad-hoc channel 10 rts 250 frag 256
```

Verifique que a porção wireless do cartão Wi-Fi foi configurada de modo a permitir uma conexão ad-hoc para outros nós mesh dentro do alcance direto (*single hop*). Certifique-se de que a interface junta-se ao mesmo canal wireless, use o mesmo nome de rede ***ESSID*** (***Extended Service Set IDentifier***) e tenha o mesmo Cell-ID que todos os outros cartões Wi-Fi usados na construção da rede mesh. Muitos cartões Wi-Fi, ou seus respectivos drivers, não são compatíveis com o padrão 802.11 para redes ad-hoc e podem falhar miseravelmente na conexão com uma célula. Eles podem ser incapazes de conectar com outros dispositivos na mesma tabela, mesmo que estejam configurados com o canal e nome de rede corretos. Eles podem até confundir outros cartões que comportam-se de acordo com o padrão ao criar seu próprio Cell-ID no mesmo canal, como o mesmo nome de rede. Cartões Wi-Fi feitos pela Intel que são embarcados com os notebooks Centrino são notórios por este tipo de comportamento.

Você pode verificar isto com o comando **`iwconfig`**, quando usar o GNU/Linux. Aqui está o resultado em minha máquina:

```
wlan0 IEEE 802.11b  ESSID:"olsr.org"
 Mode:Ad-Hoc  Frequency:2.457 GHz  Cell: 02:00:81:1E:48:10
 Bit Rate:2 Mb/s   Sensitivity=1/3
 Retry min limit:8   RTS thr=250 B   Fragment thr=256 B
 Encryption key:off
 Power Management:off
 Link Quality=1/70  Signal level=-92 dBm  Noise level=-100 dBm
 Rx invalid nwid:0  Rx invalid crypt:28  Rx invalid frag:0
 Tx excessive retries:98024 Invalid misc:117503 Missed beacon:0
```

É importante configurar o limite (*threshold*) do parâmetro ***Request to Send*** (***RTS***—solicitação para envio) para uma rede mesh. Existirão colisões no canal de rádio entre as transmissões de nós no mesmo canal wireless e o RTS atenuará isto. RTS/CTS adicionam um handshake[4] antes da transmissão de cada pacote, garantindo que o canal está livre. Isto adiciona uma sobrecarga, mas aumenta o desempenho no caso de nós escondidos—e nós escondidos são o padrão em uma rede mesh. Este parâmetro configura o limite do menor pacote (em bytes) para o qual o nó envia um RTS. O limite (*threshold*) RTS deve ser menor que o tamanho do pacote IP (*IP Packet*) e que o limite de fragmentação (*fragmentation threshold*)—aqui configurado para 256—de outra

4. N. do. T - *Handshake* traduz-se, literalmente, por "sacudida de mãos". É o tradicional cumprimento de apertar as mãos de alguém. Neste caso, é um "comportamento" do protocolo na transmissão de dados. Antes de enviar qualquer coisa, que deve ter o tamanho mínimo definido pelo threshold (no caso, 256 bytes), o lado que vai transmitir envia um sinal RTS, Request to Send, ou seja, pede permissão para enviar dados. Caso o canal esteja liberado, ele receberá como resposta um sinal CTS, Clear to Send, ou "liberado para enviar". Atendidas estas condições, a transmissão de dados se inicia.

forma, será desabilitado. O TCP é muito sensível a colisões, então é importante manter o RTS ligado.

A fragmentação permite a divisão de um pacote IP em fragmentos menores, transmitidos no meio de comunicação. Isto adiciona sobrecarga mas, em um ambiente com muito ruído, acaba por diminuir a incidência de erros e permite que os pacotes atravessem picos de interferência. Redes mesh possuem bastante ruído pois todos os seus nós usam o mesmo canal e, por causa disto, as transmissões interferem umas com as outras. Este parâmetro (*Fragment thr*) configura o tamanho máximo que um pacote deve ter, antes de ser dividido e enviado em uma rajada (*burst*). Um valor igual ao tamanho máximo do pacote IP (*IP packet size*) desabilita o mecanismo de fragmentação, desta forma, Fragment thr deve ser menor que o *IP packet size*. A configuração do limite de fragmentação é recomendada.

Uma vez que um endereço IP válido e uma máscara de rede são atribuídos, e a interface wireless está ligada, o arquivo de configuração do olsrd deve ser alterado de maneira que o olsrd encontre e use a interface na qual deve trabalhar.

Para o MAC OS X e Windows há uma boa interface gráfica para a configuração e monitoração do serviço. Infelizmente, isto é uma tentação para usuários que não têm conhecimento suficiente façam coisas estúpidas—como anunciar buracos negros. No BSD e no Linux, o arquivo de configuração /etc/olsrd.conf precisa ser manipulado com um editor de textos.

Um simples olsrd.conf

Não seria muito prático fornecer, aqui, um arquivo de configuração completo. Abaixo estão alguns itens essenciais que devem ser verificados.

```
UseHysteresis           no
TcRedundancy            2
MprCoverage             3
LinkQualityLevel        2
LinkQualityWinSize      20

LoadPlugin "olsrd_dyn_gw.so.0.3"
       {
            PlParam     "Interval"   "60"
            PlParam     "Ping"       "151.1.1.1"
           PlParam     "Ping"       "194.25.2.129"
       }

Interface "ath0" "wlan0" {
        Ip4Broadcast 255.255.255.255
       }
```

Há muito mais opções disponíveis no olsrd.conf, mas estas opções básicas servem como ponto de partida. Uma vez que estes passos estão completos, o olsrd pode ser iniciado com um simples comando no terminal:

```
olsrd -d 2
```

Recomendo que você execute o comando com a opção de debug -d 2, especialmente na primeira vez que o fizer. Desta forma, você pode ver o que o olsrd faz e monitorar de que maneira estão os links para seus vizinhos. Em

dispositivos embarcados, o nível de debug deve ser 0 (desligado), porque o debug aumenta bastante a carga da CPU.

A saída do comando anterior deve ser parecida com o seguinte:

```
--- 19:27:45.51 ------------------------------------------- DIJKSTRA

192.168.120.1:1.00 (one-hop)
192.168.120.3:1.00 (one-hop)

--- 19:27:45.51 ------------------------------------------------ LINKS

IP address           hyst     LQ     lost   total  NLQ    ETX
192.168.120.1    0.000  1.000  0      20     1.000  1.00
192.168.120.3    0.000  1.000  0      20     1.000  1.00

--- 19:27:45.51 ---------------------------------------- NEIGHBORS

IP address           LQ     NLQ    SYM   MPR   MPRS  will
192.168.120.1    1.000  1.000  YES   NO    YES       3
192.168.120.3    1.000  1.000  YES   NO    YES       6

--- 19:27:45.51 ---------------------------------------- TOPOLOGY

Source IP addr    Dest IP addr         LQ       ILQ     ETX
192.168.120.1    192.168.120.17   1.000  1.000  1.00
192.168.120.3    192.168.120.17   1.000  1.000  1.00
```

Usando OLSR em Ethernet e múltiplas interfaces

Não é necessário ter uma interface wireless para testar ou usar o olsrd—ainda que seja para isto que o olsrd tenha sido projetado. Ele pode ser usado em qualquer cartão de rede. Interfaces Wi-Fi não têm que operar sempre no modo ad-hoc para formar uma rede mesh quando um nó mesh tem mais do que uma interface. Para links dedicados, pode ser uma boa opção tê-los rodando em modo infra-estrutura. Muitos cartões Wi-Fi e seus drivers apresentam problemas em modo ad-hoc, mas funcionam bem no modo de infra-estrutura—porque todos esperam que ao menos isto funcione bem. O modo ad-hoc ainda não tem muitos usuários, assim, a implementação do mesmo foi feita de forma descuidada por muitos fabricantes. Com o aumento da popularidade de redes mesh, a situação dos drivers está melhorando hoje.

Plugins

Uma boa quantidade de plugins está disponível para o olsrd. Visite o site *www.olsr.org* para uma lista completa deles. Aqui apresentamos apenas um pequeno tutorial para o plugin de visualização de topologia de rede **`olsrd_dot_draw`**.

Com freqüência, é muito bom para o entendimento de uma rede mesh ter a capacidade de exibir a topologia da rede de forma gráfica. O plugin **`olsrd_dot_draw`** gera a topologia da rede em formato de pontos na porta TCP 2004. As ferramentas graphviz podem então ser usadas para desenhar os gráficos.

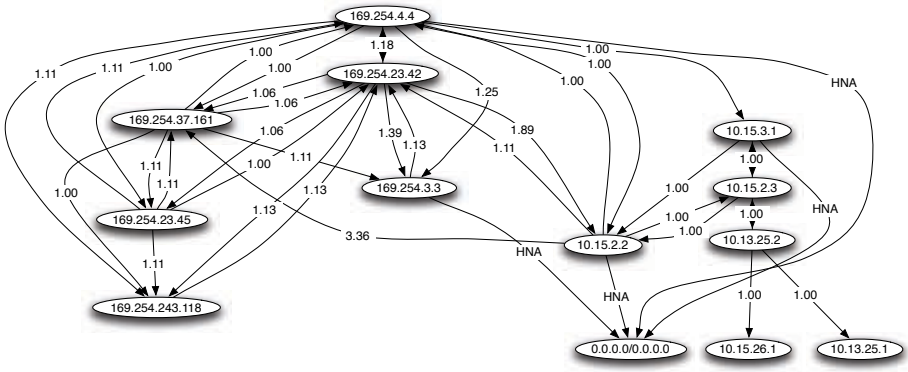


Figura 3.19: Uma topologia de rede OLSR gerada automaticamente.

Instalando o plugin dot_draw

Compile os plugins olsr separadamente e instale-os. Para carregar o plugin adicione as seguintes linhas ao **`/etc/olsrd.conf`**. O parâmetro "accept" especifica qual host é aceito para ver a informação de topologia (atualmente, apenas um) e, por padrão, é "localhost" (o computador local). O parâmetro "port" especifica a porta TCP.

```
LoadPlugin "olsrd_dot_draw.so.0.3"
       {
              PlParam "accept" "192.168.0.5"
              PlParam "port" "2004"
       }
```

Feito isto, reinicie o olsr e verifique se você obtém uma saída na porta TCP 2004.

```
telnet localhost 2004
```

Após algum tempo, você deve começar a ver alguma saída de texto.

Agora, você pode salvar a saída das descrições gráficas e rodar as ferramentas **`dot`** ou **`neato`** do pacote graphviz para obter as imagens.

Bruno Randolf escreveu um pequeno script em perl, que obtém continuamente a informação de topologia do olsrd e a exibe graficamente, utilizando as ferramentas graphviz e ImageMagick.

Comece instalando os seguintes pacotes em seu computador:

- graphviz, *http://www.graphviz.org/*
- ImageMagick, *http://www.imagemagick.org/*

Faça o download do script em *http://meshcube.org/nylon/utils/olsr-topology-view.pl*

Você pode, agora, executar o script com **`./olsr-topology-view.pl`** e verificar a atualização da topologia praticamente em tempo real.

Diagnóstico de problemas

Desde que os cartões Wi-Fi possam "ver" diretamente um ao outro com seus rádios, executar um "ping" irá funcionar, quer o olsrd esteja sendo executado ou não. Isto funciona porque a grande máscara de rede (255.255.255.255) faz de cada nó um link local, colocando de lado questões de roteamento no nível do primeiro "hop" (local). Esta é a primeira coisa a ser verificada se algo parece não estar de acordo com o esperado. A maior parte das dores de cabeça que as pessoas têm com o Wi-Fi em modo ad-hoc são causadas pelo fato de que o modo ad-hoc em cartões e drivers é implementado de forma descuidada. Se não for possível "pingar" os nós diretamente quando eles estão ao alcance um do outro, é mais provável que exista um problema no cartão ou driver, ou as configurações da sua rede estão erradas.

Se as máquinas conseguem "pingar" uma a outra, mas o olsrd não encontra as rotas, então os endereços IP, máscaras de rede e endereço de broadcast devem ser verificados.

Finalmente, você está rodando um firewall? Verifique se você não está bloqueando a porta UDP 698.

Estimando a capacidade

Links wireless podem proporcionar capacidades de transmissão de dados para seus usuários que são maiores que as conseguidas em conexões Internet tradicionais, como VSAT, linha discada, ou DSL. Esta capacidade de transmissão é também chamada de ***throughput***, ***capacidade de canal*** (***channel capacity***), ou simplesmente ***largura de banda*** (***bandwidth***)—ainda que este termo não esteja relacionado com a largura de banda de um rádio. É importante entender que a velocidade listada para um dispositivo, a taxa de transmissão de dados (***data rate***) refere-se à velocidade pela qual os rádios podem trocar símbolos, e não a capacidade utilizável que você irá observar. Como mencionado anteriormente, um simples link 802.11g pode usar rádios de 54 Mbps, mas fornecerá apenas um máximo de 22 Mbps de transmissão efetiva de dados. O restante é a sobrecarga que os rádios utilizam para coordenar seus sinais usando o protocolo 802.11g.

Note que o throughput é uma medida de bits em um determinado tempo. 22 Mbps significa que, em um segundo, até 22 megabits podem ser enviados de uma ponta de um link para a outra. Se os usuários tentarem enviar mais do que 22 megabits pelo link, isto tomará mais do que um segundo. Uma vez que os dados não possam ser enviados imediatamente, eles são colocados em uma fila (***queue***) e transmitidos tão rapidamente quanto seja possível. Este atraso no envio de dados aumenta o tempo necessário para que aqueles bits colocados mais recentemente na fila atravessem o link. Este tempo que leva para os dados atravessarem o link é chamado ***latência*** (***latency***), e uma latência alta é comumente chamada de ***lag***. Seu link irá, eventualmente, enviar todo o tráfego que está na fila, mas seus usuários provavelmente reclamarão se o *lag* for muito grande.

Quanto throughput seus usuários realmente precisam? Isto irá depender de quantos usuários você têm e como eles utilizam o link wireless. Várias aplicações Internet requerem diferentes quantidades de throughput.

Aplicação	Consumo de banda por usuário	Observações
Mensagens em texto, comunicadores instantâneos	< 1 kbps	Como o tráfego é pouco freqüente e assíncrono, programas de mensagens instantâneas toleram latências altas.
Correio eletrônico	1 a 100 kbps	Da mesma forma que programas de mensagens instantâneas, a comunicação através de email é assíncrona e, assim, tolerará latência. Grandes anexos, vírus e spam aumentam significativamente a utilização de banda. Note que serviços de webmail (como o Yahoo, Hotmail e gMail) devem ser considerados como navegação, não como email.
Navegação web	50 a mais de 100 kbps	Navegadores web apenas utilizam a rede quando dados são requisitados. A comunicação é assíncrona, assim, uma quantidade considerável de *lag* pode ser tolerada. Quando os navegadores requisitam mais dados (imagens grandes, longos downloads, etc) o uso da banda aumentará significativamente.
Streaming de áudio	96 - 160 kbps	Cada usuário de um serviço de streaming de áudio (ouvintes de rádios online, podcasts e outros) usa uma quantidade relativamente grande da largura de banda durante todo o tempo em que está ouvindo. Alguma latência pode ser tolerada com o uso de *buffers* (memória local) de bom tamanho no computador do cliente. Períodos extensos de *lag*, porém, farão com que o sinal de áudio "salte" ou que ocorram outros problemas com a sessão.
Voz sobre IP (VoIP)	24 - 100+ kbps	Assim como o streaming de áudio, o uso do VoIP compromete uma quantidade de banda de cada usuário enquanto durar a chamada. Mas com VoIP, o consumo de banda é praticamente igual em ambas as direções. A latência em uma conexão VoIP é imediata e irritante para os usuários. Uma interrupção maior que alguns milissegundos é inaceitável para VoIP.

Aplicação	Consumo de banda por usuário	Observações
Streaming de vídeo	64 - 200+ kbps	Como no streaming de áudio, alguma latência intermitente é evitada com o uso de buffers no cliente. A transmissão de vídeo requer um throughput alto e uma latência baixa para que funcione apropriadamente.
Aplicações peer-to-peer para o compartilhamento de arquivos (BitTorrent, KaZaA, Gnutella, eDonkey, etc.)	0 - infinitos Mbps	Enquanto aplicações peer-to-peer toleram qualquer quantidade de latência, elas tendem a utilizar o throughput máximo disponível para transmitir dados para quantos clientes forem possíveis. O uso destas aplicações irá causar latência e problemas de consumo de banda para todos os outros usuários da rede, a não ser que você use alguma forma cuidadosa de limitação de banda (*bandwidth shaping*)

Para estimar a largura de banda necessária para a sua rede, multiplique o número esperado de usuários pelo tipo de aplicação que eles irão, provavelmente, utilizar. Por exemplo, 50 usuários que irão, primariamente, navegar pela web consumirão entre 2,5 e 5 Mbps. Por outro lado, 50 usuários simultâneos de VoIP necessitarão de 5 Mbps ou mais de largura de banda em ***ambas as direções*** sem, absolutamente, nenhuma latência. Uma vez que os equipamentos wireless 802.1g são ***half duplex*** (ou seja, apenas transmitem ou recebem, nunca simultaneamente), você deve, apropriadamente, dobrar a largura de banda requerida para um total de 10 Mbps. Seus links wireless devem prover esta capacidade em todos os momentos, ou haverá falhas nas conversações.

Dificilmente todos os seus usuários usarão a conexão precisamente ao mesmo tempo, então, é uma prática comum superestimar (***oversubscribe***) o uso da largura de banda disponível em algum fator (isto é, permitir mais usuários que a largura de banda máxima pode suportar). Superestimar em um fator de 2 a 5 vezes o número de usuários é bastante comum. Na prática, você irá superestimar em algum fator quando estiver montando sua infra-estrutura de rede. Através do monitoramento cuidadoso do consumo de banda, você será capaz de planejar quando deve atualizar as várias partes de sua rede e quantos recursos adicionais serão necessários.

Você pode ter certeza de que, independente da capacidade que você irá fornecer, seus usuários encontrarão aplicações que irão usá-la integralmente. Conforme veremos ao final deste capítulo, o uso de técnicas de limitação de banda (*bandwidth shaping*) auxiliará na minimização de alguns problemas de latência. Com o uso de limitação de banda, armazenamento local de páginas web (*web caching*) e outras técnicas, você poderá diminuir a latência significativamente e melhorar, de maneira geral, a utilização da banda de sua rede.

Para ter uma idéia do lag percebido em conexões muito lentas, o ICTP construiu um simulador de largura de banda. Ele irá, simultaneamente, fazer a carga de uma página web em velocidade total ou a uma taxa reduzida de sua escolha. Esta demonstração dá a você um entendimento imediato de como uma taxa de transferência baixa e uma alta latência reduzem a utilidade da Internet como ferramenta de comunicação. O simulador está disponível em *http://wireless.ictp.trieste.it/simulator/*

Planejamento do link

Um sistema básico de comunicação consiste em dois rádios e suas respectivas antenas, separados por um caminho a ser coberto. Para que seja estabelecida a comunicação entre os dois rádios é necessário que as antenas captem uma certa quantidade mínima de sinal, apresentando-o ao conector de entrada do rádio. A determinação da viabilidade do link é um processo chamado "cálculo do orçamento do link" (***link budget***). A passagem, ou não, do sinal entre os rádios dependerá da qualidade do equipamento utilizado e da diminuição do sinal devido à distância (***path loss***).

Cálculo do orçamento do link

A potência disponível em um sistema 802.11 pode ser caracterizada pelos seguintes fatores:

- **Potência de transmissão** (*TX power*). É expressa em miliwatts ou dBm. Varia entre 30 mW a mais de 200 mW. A potência de transmissão freqüentemente depende da taxa de transmissão. A TX Power de um determinado dispositivo deve estar especificada na literatura fornecida pelo fabricante, mas pode ser difícil de ser encontrada às vezes. Bases de dados online, como a fornecida pela SeattleWireless (*http://www.seattlewireless.net/HardwareComparison*) podem ajudar.
- **Ganho da antena**. Antenas são componentes passivos que criam o efeito de amplificação em função de sua forma física. Elas têm as mesmas características, tanto na transmissão quanto na recepção. Assim, uma antena de 12 dBi é simplesmente uma antena de 12 dBi, sem importar se está em modo de transmissão ou recepção. Antenas parabólicas têm um ganho de 19 a 24 dBi, antenas omnidirecionais têm de 5 a 12 dBi e antenas setoriais têm um ganho aproximado de 12 a 15 dBi.
- **Nível mínimo de sinal para recepção** (*Minimum Received Signal Level*, ou RSL mínimo) expressa simplesmente a sensitividade do receptor. O mínimo RSL é sempre expresso como um dBm negativo (- dBm) e é o sinal de menor potência que o receptor consegue distinguir. O mínimo RSL depende da taxa de transmissão mas, como regra geral, a menor taxa (1 Mbps) implica na maior sensitividade. O mínimo ficará tipicamente entre -75 a -95 dBm. Como a TX Power, as especificações do RSL devem ser fornecidas pelo fabricante do equipamento.
- **Perdas em cabos**. Parte da energia do sinal é perdida nos cabos, conectores e outros dispositivos que estão entre os rádios e as antenas.

A perda depende do tipo de cabo usado e de seu comprimento. A perda de sinal em cabos coaxiais curtos, incluindo seus conectores, é bem pequena, na faixa de 2 a 3 dB. É melhor manter cada cabo sempre o mais curto possível.

Quando se calcula a perda de energia em um caminho de transmissão, vários efeitos devem ser considerados. É necessário levar em conta a ***perda em espaço aberto*** (***free space loss***), ***atenuação*** (***attenuation***) e ***espalhamento*** do sinal (***scattering***). A potência do sinal diminui com o espalhamento geométrico da frente de onda (*free space loss*). Ignorando todo o resto, quanto maior a distância entre os rádios, menor é o sinal recebido em função da perda de sinal em espaço aberto. Isto não depende do ambiente, mas apenas da distância. Esta perda acontece porque o sinal irradiado se expande em função da distância do transmissor.

Usando decibéis para expressar a perda, e usando 2,45 GHz como a freqüência do sinal, a equação que define a perda em espaço aberto é a seguinte:

$$L_{fsl} = 40 + 20*\log(r)$$

Onde L_{fsl} é expresso em dB e r é a distância entre o transmissor e o receptor, em metros.

A segunda contribuição para as perdas é dada pela atenuação. Ela acontece porque a potência do sinal é absorvida quando o mesmo atravessa objetos sólidos como árvores, paredes, janelas e separações entre andares em um prédio. A atenuação pode variar bastante, dependendo da estrutura do objeto que o sinal está atravessando e é muito difícil de ser quantificada. A maneira mais conveniente de expressar a sua contribuição para o total de perdas é adicionando uma "perda permitida" (allowed loss) para o espaço aberto. Por exemplo, a experiência mostra que árvores adicionam entre 10 a 20 dB de perda para cada uma que esteja no caminho direto do sinal, enquanto paredes contribuem de 10 a 15 dB, dependendo do tipo de material com que foram construídas.

Ao longo do caminho do link, a energia de rádio-freqüência deixa a antena de transmissão e se espalha. Uma parte da energia de RF atinge diretamente a antena de recepção, enquanto outra choca-se com o chão. Parte desta energia que chocou-se com o chão também atinge a antena de recepção. Uma vez que o sinal refletido percorre um caminho maior, ele atinge a antena depois que o sinal direto. Este efeito é chamado ***multipath*** (caminho múltiplo) ou dispersão de sinal. Em alguns casos, os sinais refletidos somam-se, sem causar problemas. Quando eles combinam-se em fases diferentes, o sinal recebido é praticamente sem valor. Em alguns casos, o sinal recebido pela antena pode ser totalmente anulado pelos sinais refletidos. Isto é conhecido como anulação (***nulling***). Há uma técnica simples que é usada para lidar com o efeito multipath, chamada de diversidade de antena (***antenna diversity***). Ela consiste na adição de uma segunda antena ao rádio. O multipath é, de fato, um fenômeno essencialmente dependente da localização. Se dois sinais adicionam-se fora de fase em um local, eles não irão adicionar-se de forma igualmente destrutiva em outro local próximo. Com duas antenas, ao menos uma dela deve ser capaz de receber um sinal utilizável, mesmo que a outra receba um distorcido. Em

dispositivos comerciais, a diversidade de troca de antenas é utilizada: há múltiplas antenas em múltiplas entradas para um único receptor. O sinal é, portanto, recebido por apenas uma antena de cada vez. Na transmissão, o rádio usa a última antena usada para a recepção. A distorção vinda do efeito multipath degrada a habilidade do receptor recuperar o sinal, de uma forma bastante similar à perda de sinal. Uma simples forma de aplicar o efeito do espalhamento no cálculo de perdas no caminho consiste na mudança do expoente do fator da distância na fórmula de perda no espaço aberto. O expoente tende a aumentar se o ambiente é propício a muito espalhamento de sinal. Um expoente 3 pode ser usado em um ambiente externo, enquanto um expoente 4 pode ser usado em um ambiente interno.

Quando as perdas no espaço aberto, a atenuação e o espalhamento são combinados, a perda total no caminho fica:

```
L (dB) = 40 + 10*n*log (r) + L (permitido)
```

Para uma estimativa rápida da viabilidade de um link, pode-se avaliar apenas a perda no espaço aberto. O ambiente pode adicionar perdas posteriores de sinal e deve ser considerado para uma avaliação exata do link. O ambiente é, de fato, um fator muito importante, que jamais deve ser negligenciado.

Para avaliar se um link é viável, as características do equipamento usado devem ser conhecidas e as perdas no caminho devem ser avaliadas. Note que, ao fazer estes cálculos, você deve considerar apenas a potência de transmissão de um lado do link. Se você usar rádios diferentes em cada lado do link, deve calcular a perda no caminho duas vezes, uma para cada direção (usando a potência de TX apropriada em cada cálculo). Adicionando todos os ganhos e subtraindo todas as perdas, temos:

```
  TX Power Rádio 1
+ Ganho de Antena Rádio 1
- Perdas no Cabo Rádio 1
+ Ganho de Antena Rádio 2
- Perdas de Cabo Rádio 2
  ————————————————————————
=       Ganho Total
```

Subtraindo a perda que ocorre no caminho do ganho total:

```
  Ganho Total
- Perda no Caminho
  ————————————————————————————————
= Nível de sinal de um lado do link
```

Se o sinal resultante for maior do que o nível mínimo de sinal de recepção, então o link é viável. O sinal recebido é potente o suficiente para que os rádios o utilizem. Lembre-se que o mínimo RSL é sempre expresso com um dBm negativo, assim -56 dBm é maior que -70 dBm. Em um determinado caminho, a variação de perdas de sinal em um período de tempo pode ser grande, desta forma uma certa margem (a diferença entre o nível de sinal e o mínimo DSL) deve ser considerada. Esta margem é a quantidade de sinal acima da sensitividade do rádio que irá garantir um link de rádio estável e de alta qualidade durante uma tempestade ou outras perturbações atmosféricas. Uma margem de 10 a 15 dB já é boa. Para dar algum espaço para a atenuação pelo

efeito multipath no sinal recebido, uma margem de 20 dB deve ser suficientemente segura.

Agora que você calculou o orçamento do link em uma direção, repita o cálculo para outra. Substitua a potência de transmissão do segundo rádio e compare o resultado com o nível de sinal mínimo de recepção do primeiro.

Exemplo de cálculo de orçamento do link

Como exemplo, vamos estimar a viabilidade de um link de 5 km, com um access point e um rádio para o cliente. O access point está conectado a uma antena omnidirecional com um ganho de 10 dBi, enquanto o cliente está conectado a uma antena setorial com um ganho de 14 dBi. A potência de transmissão do AP é 100 mW (ou 20 dBm) e sua sensitividade é de -89 dBm. A potência de transmissão do cliente é de 30 mW (ou 15 dBm) e sua sensitividade é de -82 dBm. Os cabos são curtos, com uma perda de 2 dB em cada lado.

Adicionando todos os ganhos e subtraindo todas as perdas do AP ao cliente temos:

```
  20 dBm (TX Power Rádio 1)
+ 10 dBi (Ganho de antena Rádio 1)
-  2 dB  (Perdas no cabo Rádio 1)
+ 14 dBi (Ganho de antena Rádio 2)
-  2 dB  (Perdas no cabo Rádio 2)
_____________________________________
= 40 dB Ganho Total
```

A perda no caminho, para um link de 5 km, considerando apenas a perda no espaço aberto é:

```
Perda no caminho = 40 + 20 log (5000) = 113 dB
```

Subtraindo a perda no caminho do ganho total:

```
40 dB - 113 dB = -73 dBm
```

Uma vez que -73 dB é maior que a sensitividade mínima do rádio cliente (-82 dBm), o nível do sinal está justo o suficiente para que o cliente "ouça" o access point. Há apenas uma margem de 9 dB (82 dB – 73 dB), o que significa que o link provavelmente funcionará bem com tempo bom, mas que pode não ser o bastante em más condições do tempo.

A seguir, vamos calcular o link do cliente de volta ao access point:

```
  15 dBm (TX Power Rádio 2)
+ 14 dBi (Ganho de antena Rádio 2)
-  2 dB  (Perdas no cabo Rádio 2)
+ 10 dBi (Ganho de antena Rádio 1)
-  2 dB  (Perdas no cabo Rádio 1)
_____________________________________
  35 dB = Ganho Total
```

Obviamente, a perda no caminho é a mesma na volta. Então, o sinal que recebemos do lado do access point é:

```
35 dB - 113 dB =  -78 dBm
```

Como a sensitividade do AP é -89 dBm, isto nos deixa uma pequena margem de 11 dB (89dB – 78dB). De maneira geral, este link provavelmente funcionará, mas poderia ter um pouco mais de ganho. O uso de um receptor do tipo prato no cliente, ao invés de uma antena setorial de 14 dBi, proporcionará um ganho adicional de 10 dBi em ambas as direções do link (lembre-se, o ganho da antena é recíproco). Uma opção mais cara seria usar rádios mais potentes em ambos os lados do link, mas note que a adição de um amplificador ou um cartão que proporcione maior potência apenas de um lado não ajudará, normalmente, a qualidade global do link.

Ferramentas online podem ser usadas para o orçamento do link. Por exemplo, a *Green Bay Professional Packet Radio's Wireless Network Link Analysis* (*http://my.athenet.net/~multiplx/cgi-bin/wireless.main.cgi*) é uma excelente ferramenta. A Super Edition gera um arquivo PDF contendo a zona Fresnel e os gráficos de caminhos do rádio. Os scripts para os cálculos podem até ser baixados do site e instalados localmente.

O site Terabeam também dispõe de calculadoras excelentes (*http://www.terabeam.com/support/calculations/index.php*).

Tabelas para o orçamento do link

Para calcular o orçamento do link, simplesmente use a distância aproximada de seu link e a preencha nas tabelas a seguir:

Perda no espaço livre a 2,4 GHz

Distância (m)	100	500	1.000	3.000	5.000	10.000
Perda (dB)	80	94	100	110	113	120

Para mais distâncias em perdas no caminho, consulte o **Apêndice C**.

Ganho de Antena:

Antena do Rádio 1 (dBi)	+ Antena do Rádio 2 (dBi)	= Ganho total de antena

Perdas:

Perda no cabo para o Rádio 1 (dB)	+ Perda no cabo para o Rádio 2 (dB)	+ Perda no espaço aberto (dB)	= Perda total (dB)

Orçamento do link do Rádio 1 para o Rádio 2:

TX Power do Rádio 1	+ Ganho de antena	- Perda total	= Sinal	> Sensitivitade do Rádio 2

Orçamento do link do Rádio 2 para o Rádio 1:

TX Power do Rádio 2	+ Ganho de antena	- Perda total	= Sinal	> Sensitivitade do Rádio 1

Caso o sinal recebido seja maior que o mínimo nível de sinal aceitável em ambas as direções do link, considerando os ruídos recebidos ao longo do caminho, então o link é possível.

Softwares para planejamento do link

Mesmo que o cálculo manual do orçamento do link seja simples, há uma série de ferramentas que podem ajudar na automação deste processo. Adicionalmente ao cálculo da perda no espaço aberto, estas ferramentas levam em conta muitos outros fatores relevantes (como a absorção por árvores, efeitos do terreno, clima e mesmo a estimativa de perda de sinal em áreas urbanas). Nesta sessão, discutiremos duas ferramentas livres que são úteis para o planejamento de links wireless: os utilitários online para projeto de rede do *Green Bay Professional Packet Radio* e o *Radio Mobile*.

CGIs para projeto interativo

O grupo *Green Bay Professional Packet Radio* (GBPRR) disponibiliza, online e gratuitamente, uma variedade de ferramentas úteis para o planejamento de links. Você pode acessá-las em http://www.qsl.net/n9zia/wireless/page09.html. Elas funcionarão em qualquer dispositivo que tenha um navegador web e uma conexão com a Internet.

Vamos olhar com detalhe a primeira ferramenta: ***Wireless Network Link Analysis*** (Análise de link de rede wireless), que pode ser encontrada em *http://my.athenet.net/~multiplx/cgi-bin/wireless.main.cgi*.

Para começar, digite o canal a ser usado no link. Ele pode ser especificado em MHz ou GHz. Caso você não saiba a freqüência, consulte a tabela no **Apêndice B**. Note que a tabela lista a freqüência central do canal, enquanto a ferramenta pede que seja digitada a maior freqüência transmitida. A diferença no resultado final é mínima, assim, sinta-se a vontade para usar a freqüência central. Mas, para usar a freqüência maior do canal, basta adicionar 11 MHz à freqüência central.

A seguir, digite os detalhes para o transmissor do outro lado do link, incluindo o tipo de linha de transmissão, ganho de antena e outros detalhes. Procure preencher todos os dados que conhecer ou puder estimar. Você pode até colocar a altura da antena e a elevação do local onde ela está montada. Estes dados serão usados para o cálculo do ângulo de balanço (*tilt angle*) da antena. Para o cálculo de passagem da zona Fresnel, você precisará da ferramenta *Fresnel Zone Calculator* do GBPRR.

A próxima sessão é bastante similar, mas inclui informação sobre a outra extremidade do link. Preencha os dados disponíveis nos campos apropriados.

Finalmente, a última sessão descreve o clima, o terreno e a distância do link. Preencha com tantos dados que souber ou puder estimar. A distância do link pode ser calculada fornecendo a latitude e longitude de ambas as extremidades do link, ou digitada manualmente.

Feito isto, clique no botão Submit para obter um relatório detalhado acerca do link proposto. Ele irá incluir todos os dados digitados, assim como a estimativa de perdas no caminho, taxas de erro e disponibilidade do link. Todos estes números são teóricos, mas podem dar uma idéia básica da viabilidade do link. Através do ajuste de valores no formulário, você pode fazer um exercício do tipo "e se...", verificando como a mudança de determinados parâmetros irão afetar a conexão.

Adicionalmente a ferramenta de análise básica do link, o GBPRR fornece uma "*super edition*", que produz um relatório no formato PDF, assim como uma série de outras ferramentas úteis (incluindo uma para o cálculo da zona Fresnel, um conversor de decibéis e uma calculadora para distância e direção, para ficar apenas em algumas delas). O código-fonte para a maioria das ferramentas também é fornecido.

Radio Mobile

Radio Mobile é uma ferramenta para o projeto e simulação de sistemas wireless. Ele prevê o desempenho de um link de rádio através do uso de informações sobre o equipamento e um mapa digital da área a ser coberta. O software é de domínio público, para o Windows, que pode também ser usado no Linux com o uso do emulador Wine.

O Radio Mobile usa um modelo digital de elevação de terreno (***digital terrain elevation model***) para o cálculo de cobertura, indicando a potência do sinal recebido em vários pontos ao longo do caminho. Ele automaticamente constrói um perfil entre dois pontos no mapa digital, mostrando a área de cobertura e a primeira zona Fresnel. Durante a simulação, ele verifica a linha de visão e calcula a perda no caminho, incluindo a que é devido a obstáculos. É possível criar redes de diferentes topologias, incluindo *master/slave*, ponto-a-ponto e multiponto. Ele funciona para sistemas que utilizam freqüências entre 100 kHz e 200 GHz. Mapas digitais de elevação (***Digital elevation maps*** – ***DEM***) estão livremente disponíveis a partir de várias fontes, cobrindo a maior parte do mundo. Os DEMs não mostram linhas costeiras ou outros acidentes geográficos facilmente identificáveis, mas podem ser facilmente combinados com outras fontes de dados (como fotografias aéreas ou mapas topográficos) em múltiplas camadas, para que seja possível a obtenção de representações geográficas de maior utilidade e representação mais fácil. Você também pode

digitalizar seus próprios mapas e combiná-los com DEMs. Mapas de elevação digital podem ser combinados com mapas digitalizados, fotos de satélite e serviços de mapas disponíveis na Internet (como o Google Maps) para a produção de projetos com bastante precisão.

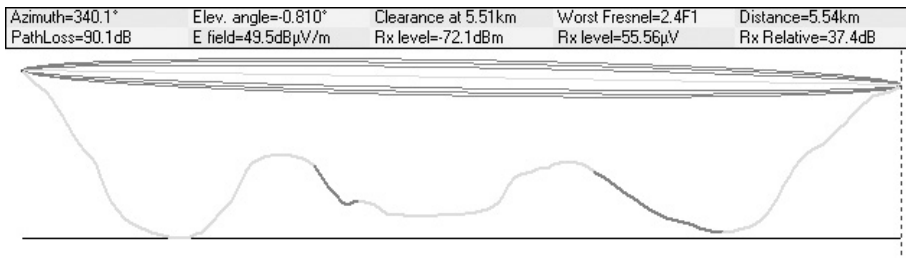


Figura 3.20: Viabilidade de link, incluindo estimativas de zona Fresnel e linha de visão, utilizando o Radio Mobile.

A página principal do Radio Mobile, com exemplos e tutoriais, está disponível em *http://www.cplus.org/rmw/english1.htm*

Radio Mobile com Linux

É possível utilizar o Radio Mobile com o Ubuntu Linux mas, enquanto a aplicação é executada, algumas legendas de botões podem ficar escondidas sobre a moldura dos mesmos e serem de difícil leitura.

Conseguimos executar o Radio Mobile em uma máquina Linux com o seguinte ambiente[5]:

- IBM Thinkpad x 31
- Ubuntu Breezy (v5.10), *http://www.ubuntu.com/*
- Wine versão 20050725, do repositório Ubuntu Universe

Há instruções detalhadas para a instalação do Radio Mobile no Linux em *http://www.cplus.org/rmw/english1.htm*l. Você deve seguir todos os passos listados, exceto o primeiro (uma vez que é difícil extrair uma DLL do arquivo VBRUN60SP6.EXE[6] no Linux). Você precisará de uma cópia do MSVBVM60.DLL de uma máquina Windows onde o run-time do Visual Basic 6 esteja instalado, ou simplesmente busque no Google o arquivo MSVBVM60.DLL e baixe-o para a sua máquina.

Agora, continue com o segundo passo do site acima, certificando-se de que a descompactação dos arquivos seja feita no mesmo diretório em que você colocou o DLL. Note que você não precisará preocupar-se com os passos

5. N. do T. - O tradutor conseguiu executar o Radio Mobile em uma distribuição Linux Mint Elyssa (baseada no Ubuntu Hardy), usando o Wine versão 0.9.59. Provavelmente, qualquer distribuição que aceite esta, ou qualquer outra versão mais recente do Wine, deve servir ao propósito. Os problemas com as legendas dos botões, descritos no texto, não foram sentidos.

6. N. do T. - Apenas usando o Wine, o tradutor pôde instalar apropriadamente o arquivo DLL, usando as instruções do site original do Radio Mobile.

seguintes ao quarto, uma vez que estes são necessários apenas para os usuários do Windows.

Para executar o programa, basta clicar duas vezes sobre ele, usando seu gerenciador de arquivos, ou através de um terminal com o seguinte comando:

```
# wine RMWDLX.exe
```

Isto deve apresentar a janela inicial do Radio Mobile em sua interface gráfica.

Evitando ruídos

As bandas irrestritas ISM e U-NII representam uma peça muito pequena do espectro eletromagnético conhecido. Como esta região pode ser usada sem o pagamento de licenças, muitos dispositivos a utilizam para um amplo número de aplicações. Telefones sem fio, transmissores analógicos de vídeo, Bluetooth, monitores para bebês e mesmo fornos de microondas competem com redes wireless no uso da bastante limitada banda de 2,4 GHz. Os sinais destes dispositivos, em conjunto com o sinal de outras redes wireless, podem causar problemas significativos para redes sem fio de longo alcance. Aqui apresentamos alguns passos que você pode seguir para reduzir a recepção de sinais indesejados.

- **Aumente o ganho da antena em ambos os lados de um link ponto-a-ponto**. Antenas não apenas adicionam ganho a um link, mas também aumentam sua direcionalidade e tendem a rejeitar ruídos vindos de áreas nos arredores do mesmo. Duas antenas parabólicas (em forma de prato) que apontam uma para a outra irão rejeitar sinais que estão fora do caminho do link. Antenas omnidirecionais receberão ruído de todas as direções.
- **Use antenas setoriais ao invés de usar uma antena omnidirecional**. Com o uso de múltiplas antenas setoriais, você pode reduzir a quantidade de ruído recebida em um ponto de distribuição. Com a divisão dos canais usados em cada setor, você pode também aumentar a largura de banda disponibilizada para seus clientes.
- **Não use um amplificador**. Como você verá no **Capítulo 4**, estes dispositivos podem piorar a recepção por amplificar indiscriminadamente todos os sinais recebidos, incluindo os de fontes de interferência. Os amplificadores também podem ser a causa de problemas de interferência para os outros usuários nas vizinhanças da banda utilizada.
- **Use o melhor canal disponível**. Lembre-se que os canais 802.11b/g têm a largura de 22 MHz, mas são separados por apenas 5 MHz. Faça uma pesquisa nos locais onde instalará seus equipamentos e escolha um canal que esteja o mais longe possível de fontes de interferência existentes. Lembre-se que o cenário wireless pode mudar a qualquer momento, uma vez que as pessoas podem passar a usar novos dispositivos (telefones sem fio, outras redes, etc). Se o seu link começar, repentinamente, a ter problema de transmissão de pacotes, você talvez precise fazer uma nova pesquisa de seu ambiente e selecionar um canal diferente.

- **Use pequenos saltos (hops) e repetidores, ao invés de cobrir uma longa distância com um link único**. Mantenha seus links ponto-a-ponto tão curtos quanto possível. Mesmo que seja possível criar um link de 12 kms que passe pelo meio de uma cidade, é bem provável que você tenha muitos problemas com interferências. Se você puder dividir este link em dois ou mais saltos (*hops*) curtos, ele ganhará mais estabilidade. Obviamente, isto não é viável em links de longa distância em áreas rurais, onde não há estruturas de rede elétrica ou suportes para antenas. Mas neste caso, problemas com ruídos também são improváveis.

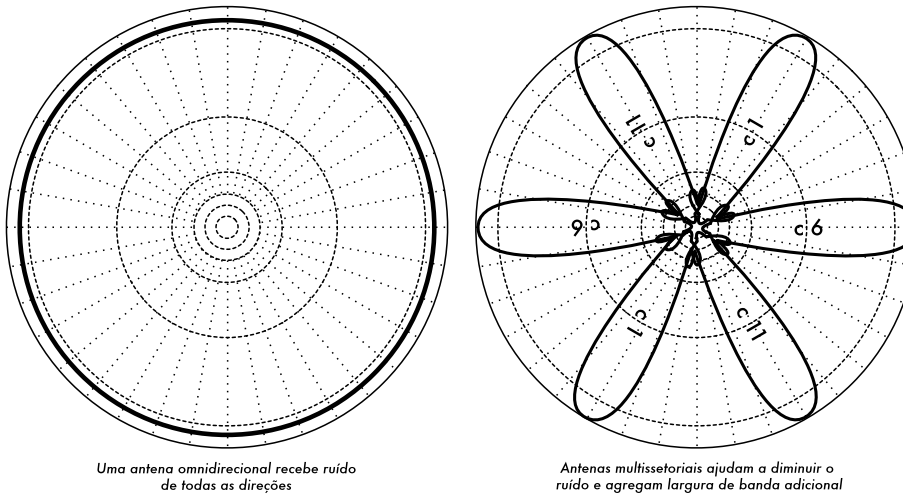


Figura 3.21: Uma antena omnidirecional em comparação com antenas multissetoriais.

- **Se possível, use bandas livres de 5,8 GHz, 900 MHz ou outras**. Mesmo sendo uma solução de curto prazo, hoje há muito mais equipamentos instalados que usam a freqüência de 2,4 GHz. Usando 802.11a ou um dispositivo que eleva a freqüência de 2,4 GHz para 5,8 GHz permitirá que o congestionamento seja evitado. Caso você consiga encontrá-los, alguns equipamentos antigos 802.11 usam o espectro livre de 900 MHz (infelizmente, com taxas de transmissão bem menores). Outras tecnologias, como Ronja (*http://ronja.twibright.com/*) usam a transmissão ótica para a implantação de links de curta distância, livres de ruído.
- **Se tudo isso falhar, use o espectro sob licença**. Existem lugares onde todo o espectro livre disponível já é, efetivamente, utilizado. Nestes casos, pode fazer sentido gastar mais dinheiro para a aquisição de equipamentos proprietários que usam bandas menos congestionadas. Para links ponto-a-ponto de longa distância que requerem uma alta taxa de transmissão esta é, certamente, uma opção. Claro que estas funcionalidades estão disponíveis em uma faixa de preço bem mais alta, comparada com a de equipamentos que operam nas bandas livres.

Para identificar as fontes de ruído, você precisará de ferramentas que mostrem o que está acontecendo no ar, em 2,4 GHz. Veremos alguns exemplos destas ferramentas no **Capítulo 6**.

Repetidores

O componente mais crítico na construção de links de longa distância é a linha de visão (***line of sight***, ou ***LOS***). Sistemas terrestres de microondas simplesmente não toleram altas colinas, árvores ou outros obstáculos no caminho de uma conexão de longa distância. Você deve ter uma boa idéia da topografia do espaço entre os dois pontos que deseja conectar, antes de determinar se o link é mesmo possível.

Mas mesmo que exista uma montanha entre dois pontos, lembre-se que os obstáculos podem, às vezes, serem usados em nosso benefício. Montanhas podem bloquear seu sinal, mas, assumindo que uma rede elétrica esteja disponível, elas podem ser locais muito bons para a instalação de repetidores.

Repetidores (***repeaters***) são nós que estão configurados para retransmitir o tráfego que não tem por destino o próprio nó. Em uma rede mesh, todo nó é um repetidor. Em uma infra-estrutura de rede tradicional, nós podem ser configurados para passar adiante o tráfego para outros nós.

Um repetidor pode usar um ou mais dispositivos wireless. Quando apenas um rádio é utilizado (repetidor de um braço só, *one-arm repeater*), a eficiência é, de maneira geral, um pouco menor que a metade da banda disponível, uma vez que o rádio pode apenas transmitir ou receber dados, nunca os dois ao mesmo tempo. Estes dispositivos são mais baratos, mais simples e possuem requerimentos menores de energia elétrica. Um repetidor com dois ou mais rádios pode operá-los em sua capacidade total, desde que eles estejam configurados para que não usem canais que interfiram um com o outro. Claro que repetidores podem também prover uma conexão Ethernet para equipamentos locais.

Repetidores podem ser adquiridos como uma solução completa de hardware ou implementados de forma simples através da conexão de um ou mais nós wireless por meio de um cabo Ethernet. Quando planejar o uso de um repetidor construído com a tecnologia 802.11, lembre-se que os nós devem ser configurados para o modo master, gerenciado ou ad-hoc. Tipicamente, ambos os rádios em um repetidor são configurados em modo master, permitindo que múltiplos clientes conectem-se com qualquer lado do repetidor. Mas dependendo do projeto de sua rede, um ou mais dispositivos podem necessitar o uso do modo ad-hoc ou mesmo do modo cliente.

Tipicamente, repetidores são usados para superar obstáculos no caminho de um link de longa distância. Por exemplo, podem existir prédios em seu caminho, mas nestes prédios moram pessoas. Com freqüência, podem ser feitos acordos com o proprietário do prédio para que a oferta de serviços de conectividade a eles seja trocada pelos direitos de uso do telhado e da eletricidade. Mesmo que o proprietário do prédio não esteja interessado, os moradores dos andares mais altos podem ser persuadidos a instalar o equipamento em uma janela.

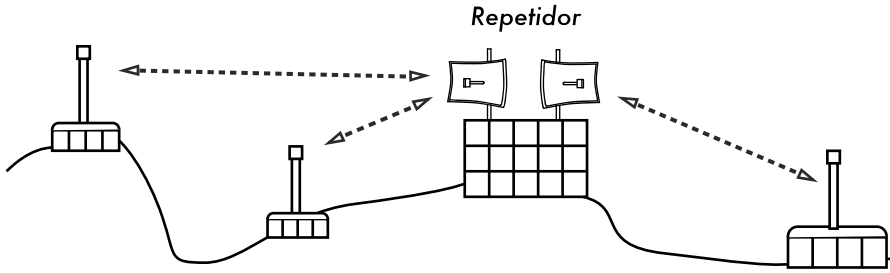


Figura 3.22: O repetidor encaminha pacotes entre os nós que não possuem uma linha de visão direta entre eles.

Caso você não possa atravessar um obstáculo, você pode contorná-lo. Ao invés de usar um link direto, tente um projeto com múltiplos hops para evitá-lo.

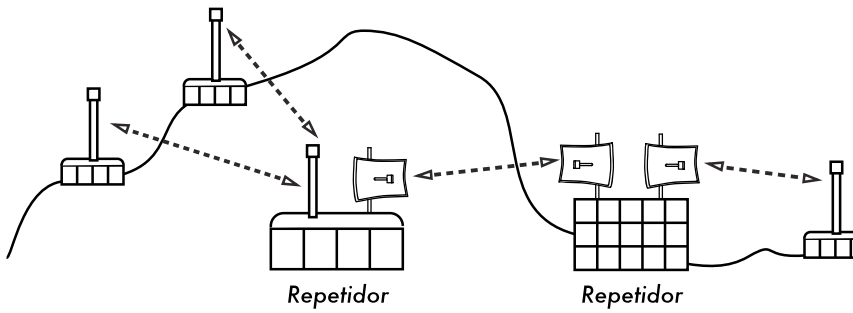


Figura 3.23: Não havia energia elétrica no topo da colina, mas a mesma foi contornada com o uso de múltiplos repetidores ao redor da base.

Finalmente, você deve considerar ir para trás, ao invés de ir adiante. Caso exista um local alto disponível em uma direção diferente, mas que possa ser visto além do obstáculo, uma conexão estável poderá ser possível com o uso de uma rota indireta.

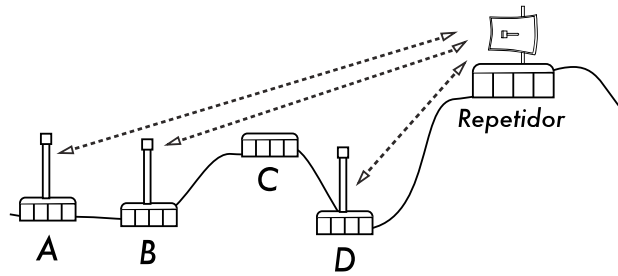


Figura 3.24: A localidade D não consegue estabelecer uma conexão direta para a localidade A ou B, pois C está no caminho e não há nenhum nó disponível ali. Com a instalação de um repetidor alto, os nós A, B e D podem comunicar-se. Note que o tráfego do nó D primeiro afasta-se da rede, antes que o repetidor encaminhe-o adequadamente.

Os repetidores em redes fazem-me lembrar do princípio de “seis graus de separação”. A idéia é a de que não importa quem você está procurando, você precisará apenas contatar cinco intermediários antes de encontrar tal pessoa. Repetidores em localidades altas podem “ver” um grande número de intermediários, e desde que seu nó esteja ao alcance do repetidor, você pode comunicar-se com qualquer outro nó que o repetidor possa alcançar.

Otimização de tráfego

A largura de banda é a medida da quantidade de bits que podem ser transmitidos em um intervalo de tempo. Isto quer dizer que, na medida que o tempo passa, a largura de banda disponível em um link aproxima-se do infinito. Infelizmente, para qualquer período definido de tempo, a largura de banda fornecida por qualquer rede não é infinita. Você sempre pode fazer o download (ou upload) da quantidade de informação que quiser; apenas terá que esperar o tempo suficiente para isso. Claro que usuários humanos não são tão pacientes quanto os computadores e não desejam esperar um tempo infinito para que seus dados atravessem a rede. Por esta razão, a largura de banda deve ser gerenciada e priorizada da mesma forma que qualquer outro recurso limitado.

Você aumentará significativamente o tempo de resposta e maximizará a taxa de transferência disponível através da eliminação do tráfego não desejado e redundante em sua rede. Esta sessão descreve algumas técnicas comuns que garantem que sua rede apenas carregue o tráfego que deve atravessá-la. Para uma discussão mais aprofundada do complexo assunto de otimização de ocupação de banda, leia o livro *How to Accelerate Your Internet* (*http://bwmo.net/*), que está disponível livremente.

Web caching

Um web proxy é um servidor em sua rede local que mantém cópias de páginas recentemente vistas, ou as mais freqüentemente visitadas, ou partes de páginas na web. Quando a próxima pessoa buscar estas páginas, elas serão servidas pelo servidor proxy local ao invés de virem da Internet. Isto resulta, na maioria dos casos, em um acesso web mais rápido, ao mesmo tempo em que reduz, de forma geral, o uso da largura de banda na Internet. Quando um servidor proxy é implementado, o administrador deve também saber que algumas páginas não podem ser armazenadas localmente – por exemplo, aquelas geradas por scripts no servidor, ou outros conteúdos gerados dinamicamente.

A aparente carga das páginas web também é afetada. Com uma conexão de baixa velocidade com a Internet, uma página estática começa a ser carregada lentamente, primeiro mostrando algum texto e depois as figuras, uma a uma. Em uma rede com um servidor proxy, pode haver a percepção de um pequeno atraso, dentro do qual nada parece acontecer, e então a página é carregada quase imediatamente. Isto acontece porque a informação é enviada tão rapidamente ao computador que ele necessita de uma quantidade perceptível de tempo para renderizar (montar e exibir) a página. O tempo total que uma página inteira leva para ser carregada pode ser de apenas dez segundos (enquanto, sem um servidor proxy, poderia demorar 30 segundos para carregar gradualmente a página). Mas, a não ser que isto seja explicado para alguns

usuários impacientes, eles podem achar que o servidor proxy tornou as coisas mais lentas. Usualmente, é tarefa do administrador de rede lidar com questões de percepção como esta.

Produtos para servidores proxy

Há uma série de servidores web proxy disponíveis. Estes são os pacotes de software mais comumente usados:

- **Squid**. O Squid, de código aberto, é o padrão de fato em universidades. Ele é gratuito, confiável, fácil de usar e pode ser melhorado (por exemplo, com a adição de filtragem de conteúdo e bloqueio de propagandas). O Squid gera registros que podem ser analisados por programas como o Awstats ou Webalizer, ambos de código aberto, capazes de produzir bons relatórios gráficos. Na maioria dos casos, é mais fácil instalá-lo como parte de sua distribuição do que fazer o download diretamente de *http://www.squid-cache.org/* (a maioria das distribuições Linux, como o Debian, assim como outras versões de Unix como o NetBSD e o FreeBSD já têm o Squid). Um bom guia de configuração do Squid pode ser encontrado no *Squid Users Guide Wiki* em *http://www.deckle.co.za/squid-users-guide/*.
- **Microsoft Proxy server 2.0**. Não está disponível para novas instalações, pois foi sucedido pelo servidor Microsoft ISA, e não é mais suportado. Mesmo assim, ele é utilizado por algumas instituições, mas não deve ser considerado para novas instalações.
- **Microsoft ISA server**. O ISA server é um servidor proxy muito bom, mas talvez muito caro para o que faz. Entretanto, com descontos acadêmicos ele pode ser acessível para algumas instituições. Ele produz seus próprios relatórios gráficos, mas seus registros podem também ser analisados por ferramentas populares como Sawmill (*http://www.sawmill.net/*). Os administradores em localidades que usam o MS ISA Server devem dispender tempo suficiente para configurá-lo corretamente, pois, de outra forma, o próprio servidor pode tornar-se um considerável usuário de largura de banda. Por exemplo, uma instalação padrão pode facilmente consumir mais banda que a usada anteriormente, pois páginas populares com datas de expiração curtas (tais como sites de notícias) serão continuamente atualizadas. Por isso, é importante ter os parâmetros de pre-fetching (busca antecipada) definidos corretamente, e que o processo de pre-fetching ocorra, preferencialmente, na madrugada. O ISA Server pode também ser usado com produtos de filtragem de conteúdo, como o WebSense. Para mais informações, visite *http://www.microsoft.com/isaserver/* e *http://www.isaserver.org/* .

Evitando que os usuários contornem o servidor proxy

Mesmo que o contorno a censuras impostas à Internet e a política de acesso à informações restritas possam ser um louvável esforço político, proxies e firewalls são ferramentas necessárias em áreas com largura de banda

extremamente limitada. Sem tais ferramentas, a estabilidade e usabilidade da rede podem ser ameaçadas pelos próprios usuários legítimos da rede. Técnicas para contornar (*bypass*) um servidor proxy podem ser encontradas em *http://www.antiproxy.com/*. Este site é útil para administradores de rede, para que eles vejam como seus sistemas resistem a estas técnicas.

Para forçar o uso do proxy que armazena o conteúdo localmente (*caching proxy*), você deve considerar a simples elaboração de uma política de acesso à rede e confiar que seus usuários a respeitarão. Na configuração abaixo, o administrador tem que confiar que seus usuários não contornarão o servidor proxy.

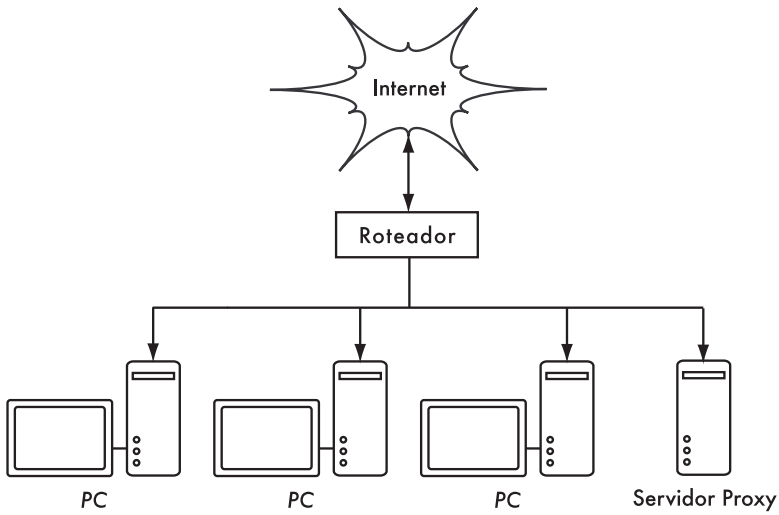


Figura 3.25: Esta rede parte do princípio de que usuários confiáveis configurem propriamente seus Pcs para que usem o servidor proxy.

Neste caso, o administrador utilizará, tipicamente, uma das seguintes técnicas:

- **Não disponibilizar o endereço do gateway padrão através de DHCP**. Isto pode funcionar por um tempo, mas alguns usuários conhecedores de redes que querem contornar o proxy podem descobrir, ou tentar adivinhar, o endereço do gateway padrão. Uma vez que tal endereço é descoberto, a tendência é que a informação sobre como contornar o proxy seja espalhada.
- **Uso de políticas de domínios ou grupos**. Isto é bastante útil para a configuração dos parâmetros de servidor proxy para o Internet Explorer em todos os computadores em um domínio, mas não garante que o proxy não seja contornado porque depende da autenticação do usuário no domínio NT. Um usuário com o Windows 95/98/ME pode cancelar sua autenticação (logon) e então contornar o proxy, e alguém que conheça a senha local em um Windows NT/2000/XP pode autenticar-se localmente e fazer o mesmo.
- **Implorar e brigar com os usuários**. Esta técnica, mesmo comum, nunca é a melhor situação para um administrador de rede.

Firewall

Uma forma mais confiável de garantir que os PCs não evitem o proxy pode ser implementada usando o firewall. O firewall pode ser configurado para permitir que apenas o servidor proxy faça solicitações HTTP para a Internet. Todos os demais PCs serão bloqueados, como mostra a **Figura 3.26**.

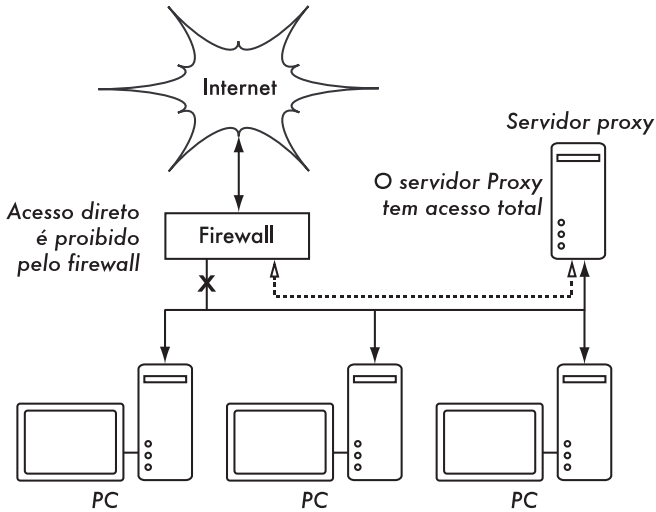


Figura 3.26: O firewall evita que os PCs acessem diretamente a Internet, mas permite este acesso através do servidor proxy.

Confiar em um firewall pode, ou não, ser suficiente, dependendo da forma como ele está configurado. Se ele apenas bloquear o acesso da rede local do campus para a porta 80 dos servidores web, usuários inteligentes ainda encontrarão meios de contornar este bloqueio. Eles também poderão usar outros serviços famintos por banda, como o BitTorrent ou Kazaa.

Dois cartões de rede

Talvez o mais confiável método seja a instalação de dois cartões de rede no servidor proxy, conectando a rede do campus à Internet como mostrado abaixo. Desta forma, a configuração da rede torna fisicamente impossível o acesso à Internet sem que se passe pelo servidor proxy.

O servidor proxy neste diagrama não deve ter o encaminhamento de IP (*IP forwarding*) habilitado, a não ser que os administradores de rede saibam exatamente o que eles querem deixar passar.

Uma grande vantagem desta configuração é que uma técnica conhecida como ***proxy transparente*** (***transparent proxying***) pode ser usada. Isto significa que os pedidos de acesso à web pelos usuários são automaticamente passados para o servidor proxy, sem nenhuma necessidade de configuração dos navegadores. A técnica faz com que, efetivamente, todo o tráfego seja armazenado localmente pelo proxy, eliminando muitas chances de erros de usuários, permitindo mesmo que se trabalhe com dispositivos que não suportem

a configuração manual para o acesso através de um proxy. Para mais detalhes sobre a configuração de um proxy transparente com o Squid, veja:

- *http://www.squid-cache.org/Doc/FAQ/FAQ-17.html*
- *http://en.tldp.org/HOWTO/mini/TransparentProxy-2.html*

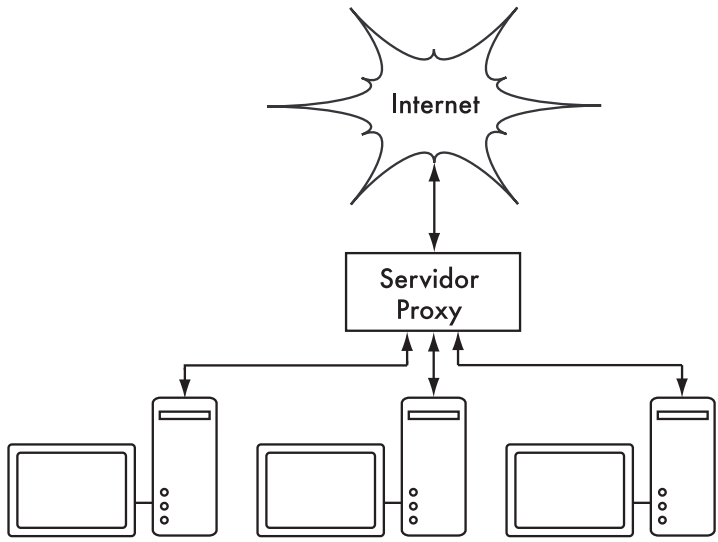


Figura 3.27: A única rota para a Internet é através do proxy.

Políticas de roteamento

Uma forma de prevenir o contorno do proxy, usando equipamento Cisco, é o uso de políticas de roteamento. O roteador Cisco direciona, de forma transparente, todo o tráfego web para o servidor proxy. Esta técnica é usada na Makerere University. A vantagem deste método é que, caso o servidor proxy não esteja funcionando, a política de roteamento pode ser temporariamente removida, permitindo que os clientes conectem-se diretamente à Internet.

Espelhando um website

Com a permissão do proprietário ou web master de um site, ele pode ser completamente espelhado para o servidor local durante a madrugada, caso não seja muito grande. Isto é algo que deve ser considerado para websites que são de particular interesse para a organização, ou aqueles que são muito populares entre os usuários. Assim como pode ser de grande utilidade, esta técnica também tem suas falhas. Por exemplo, caso o site a ser espelhado tenha scripts CGI ou outro conteúdo dinâmico que necessite de interação com o usuário, isto irá causar problemas. Uma pessoa que entra com seus dados em um site para o registro em uma conferência, onde os scripts de registro também foram espelhados localmente, pode não ter o seu registro efetivado no site real.

O espelhamento de um site pode infringir direitos autorais. Por isso, esta técnica deve apenas ser usada com a permissão formal para o site a ser

espelhado. Caso o site use ***rsync***, ele pode ser espelhado com o mesmo. Este é, provavelmente, o meio mais rápido e eficiente de manter conteúdos sincronizados. Caso o site remoto não utilize o rsync, o software recomendado é o ***wget***. Ele é parte da maioria das versões de Unix/Linux. Uma versão para Windows pode ser encontrada em *http://xoomer.virgilio.it/hherold/*, ou no pacote livre de ferramentas Cygwin Unix (*http://www.cygwin.com/*).

Um script pode ser configurado no servidor web local para que, todas as noites, faça o seguinte:

Vá para o diretório raiz de documentos do servidor web: por exemplo **`/var/www/`** no Unix, ou **`C:\Inetpub\wwwroot`** no Windows;

Faça o espelhamento do website com o seguinte comando:

```
wget --cache=off -m http://www.python.org
```

O website espelhado ficará no diretório *www.python.org*. O servidor web deve agora ser configurado para servir o conteúdo deste diretório em um host virtual. Configure o servidor DNS local para "imitar" uma entrada para este site. Para que isto funcione, os PCs clientes devem ser configurados para usar o servidor DNS local como seu DNS primário (isto é recomendável em todos os casos, já que um DNS local acelera o tempo de resposta da web).

Populе antecipadamente o cache usando wget

Ao invés de configurar o espelhamento de um website, como descrito na sessão anterior, uma técnica melhor é a de popular o proxy cache usando um processo automatizado. Este método foi descrito por J. J. Eksteen e J. P. L. Cloete do CSIR em Pretória, África do Sul, no artigo *Enhancing International World Wide Web Access in Mozambique Through the Use of Mirroring and Caching Proxies* (Melhorando o acesso internacional para a web em Moçambique através do uso de espelhamento e proxies de armazenamento local). Neste artigo (disponível em *http://www.isoc.org/inet97/ans97/cloet.htm*) eles descrevem como este processo funciona:

> *"Um processo automático busca a página principal do site e um número especificado de páginas adicionais (recursivamente seguindo os links HTML nas páginas buscadas) através do uso de um proxy. Ao invés de gravar as páginas buscadas no disco local, o processo de espelhamento as descarta. Isto é feito para conservar recursos do sistema e também para evitar possíveis conflitos de direitos autorais. Através do uso do proxy como um intermediário, as páginas buscadas ficam, com certeza, no cache do proxy, como se um usuário tivesse acessado tais páginas. Quando um cliente acessa uma página já armazenada, ela é servida do cache e não do congestionado link internacional. Este processo pode ser executado fora dos horários de pico a fim de maximizar o uso da banda e não competir com outras atividades que requeiram acesso à web."*

O comando a seguir (programado para ser executado todas as noites, ou uma vez por semana) é tudo o que é necessário (repetido para cada site que deva ser populado antecipadamente).

```
wget --proxy-on --cache=off --delete after -m http://www.python.org
```

Estas opções habilitam o seguinte:

- **-m**: Espelha o site inteiro. O wget inicia em *www.python.org* e segue os links, baixando, então, todas as páginas;
- **--proxy-on**: Espelha o site inteiro. O wget inicia em www.python.org e segue os links, baixando, então, todas as páginas;
- **--c cache-off**: Garante que o conteúdo recente seja buscado da Internet e não do servidor proxy local;
- **--delete after**: Apaga a cópia espelhada. O conteúdo espelhado é mantido no cache do proxy, desde que exista espaço suficiente em disco e os parâmetros de configuração estejam corretamente ajustados.

Adicionalmente, o wget tem muitas outras opções, por exemplo, fornecendo uma senha para sites que necessitem de uma. Quando esta ferramente é usada, o Squid deve ser configurado com o espaço em disco suficiente para conter todos os sites pré-populados e mais (para o uso normal do Squid, incluindo as demais páginas que não são pré-populadas). Felizmente, espaço em disco está se tornando cada vez mais barato e em quantidades cada vez maiores. Ainda assim, esta técnica deve ser usada apenas para alguns poucos sites selecionados. Eles não podem ser grandes a ponto de fazer com que o processo demore além da madrugada e a utilização do espaço em disco deve ser monitorada.

Hierarquias de cache

Quando uma organização tem mais de um servidor proxy, eles podem compartilhar a informação armazenada entre eles. Por exemplo, se uma página existir no cache do servidor A, mas não no cache do servidor B, um usuário conectado através do servidor B pode obter o conteúdo armazenado no servidor A através do servidor B. Os protocolos ***ICP*** (***Inter-Cache Protocol*** – protocolo inter-cache) e ***CARP*** (***Cache Array Routing Protocol*** – protocolo de roteamento para matriz de caches) podem ser usados para o compartilhamento da informação entre caches. O CARP é considerado o melhor protocolo. O Squid tem suporte a ambos os protocolos e o servidor MS ISA suporta o CARP. Para mais informações consulte *http://squid-docs.sourceforge.net/latest/html/c2075.html*. Este compartilhamento de informações reduz o consumo de banda em organizações onde mais de um proxy é utilizado.

Especificações do proxy

Na rede de um campus de uma universidade deve existir mais de um servidor proxy, tanto por questões de desempenho quanto por redundância. Com os discos baratos, de grande capacidade, disponíveis hoje, servidores proxy poderosos podem ser construídos, com 50 GB ou mais de espaço em disco alocado para o cache. O desempenho do disco é importante, por isso os discos com interface SCSI terão a melhor performance (mas um cache baseado em discos IDE ainda é melhor do que não ter um cache). O espelhamento RAID ou de outro tipo não é recomendado.

Também é recomendável ter um disco dedicado exclusivamente ao cache. Por exemplo, um disco poderia ser usado para o cache e um outro para o

sistema operacional e para os registros (logs) do cache. O Squid foi projetado para usar o máximo de memória RAM que puder, uma vez que os dados, quando estão na RAM, são obtidos com velocidade muito maior do que quando estão no disco rígido. Para a rede de um campus, a memória RAM deve ser de 1 GB ou mais:

- Além da memória requerida pelo sistema operacional e outras aplicações, o Squid necessita de 10 MB de RAM para cada 1 GB de cache em disco. Assim, se há 50 GB alocados em espaço no disco para o cache, o Squid precisará de 500 MB de memória adicional;
- A máquina também precisará de 128 MB para o Linux e 128 MB para o Xwindows;
- Mais 256 MB devem ser adicionados para outras aplicações e a fim de que tudo seja executado devidamente. Nada melhora mais o desempenho de uma máquina do que a adição de uma grande capacidade de memória, já que isto reduz a necessidade de uso do disco rígido. A memória é milhares de vezes mais rápida do que o disco rígido. Sistemas operacionais modernos mantêm os dados freqüentemente acessados na memória, desde que haja RAM disponível o suficiente, mas usam páginas de arquivo como memória extra, caso a RAM não seja suficiente.

Cache de DNS e otimização

DNSs apenas de cache (caching-only DNS) são servidores sem autoridade para nenhum domínio, apenas armazenando os resultados das buscas solicitadas pelos clientes. Assim como um servidor proxy que armazena as páginas web populares por algum tempo, os endereços DNS são armazenados até que seu tempo de vida (***TTL*** – ***time to live***) expire. Isto reduz a quantidade de tráfego DNS em uma conexão com a Internet, uma vez que o DNS cache pode satisfazer muitas das resoluções de endereço localmente. Obviamente, os computadores clientes devem ser configurados de forma a usar o servidor de nomes em cache como seu servidor DNS. Quando todos os clientes usam este servidor como seu DNS primário, ele rapidamente terá seu cache populado com a relação entre endereços IP e nomes e, desta forma, a resolução para endereços já armazenados pode ser feita rapidamente. Servidores DNS que têm autoridade sob um domínio podem também atuar como cache para mapas de endereços de hosts resolvidos por eles.

Bind (named)

O Bind é o programa padrão de fato para o serviço de nomes na Internet. Quando o Bind está instalado e em execução, ele irá atuar como um servidor caching-only, sem a necessidade de configuração adicional. O Bind pode ser instalado a partir de um pacote, como um pacote Debian ou um RPM. A instalação a partir de um pacote é, usualmente, o método mais simples. No Debian, digite:

```
apt-get install bind9
```

Além de rodar um cache, o Bind pode também rodar na forma de servidor com autoridade para zonas de nomes, servir como escravo para zonas de nomes, implementar horizonte dividido (*split horizon*) e qualquer outra configuração possível com DNS.

dnsmasq

Um servidor caching-only alternativo é o ***dnsmasq***. Existem pacotes disponíveis para o BSD e para a maioria das distribuições Linux, ou diretamente em *http://www.thekelleys.org.uk/dnsmasq/*. A grande vantagem do dnsmasq é sua flexibilidade: ele atua tanto como um caching DNS como fonte de autoridade para servidores e domínios, sem a complexidade de um arquivo de configuração de zonas. As zonas podem ser atualizadas sem a necessidade de reinicializar o serviço. Ele também pode servir como um servidor DHCP e integrar o serviço de DNS com as solicitações de DHCP dos servidores. Ele é bastante leve, estável e extremamente flexível. O Bind é, provavelmente, a melhor escolha para redes muito grandes (mais de algumas centenas de nós), mas a simplicidade do dnsmasq o torna atrativo para redes de tamanho pequeno ou médio.

Windows NT

Para instalar o serviço de DNS em um Windows NT4: selecione *Control Panel* ➔ *Network* ➔ *Services* ➔ *Add* ➔ *Microsoft DNS server*. Coloque o CD do Windows NT4 quando solicitado. A configuração de um servidor caching-only no NT é descrita no artigo 167234 do Knowledge Base. Do artigo:

> *"Simplesmente instale o DNS e execute o Domain Name System Manager. Clique em DNS no menu, selecione New Server e digite o endereço IP do computador onde você tem o DNS instalado. Você agora tem um servidor DNS caching- only."*

Windows 2000

Instalação do serviço DNS: *Start* ➔ *Settings* ➔ *Control Panel* ➔ *Add/Remove Software*. Em *Add/Remove Windows Components*, selecione *Components* ➔ *Networking Services* ➔ *Details* ➔ *Domain Name System (DNS)*. Agora execute o DNS MMC (*Start* ➔ *Programs* ➔ *Administrative Tools* ➔ *DNS*). No menu Action selecione "*Connect To Computer*...". Na janela *Select Target Computer*, habilite "*The following computer*:" e coloque o nome do servidor DNS do qual você quer fazer cache. Se houver um . [ponto] no DNS manager (isto aparece como padrão), significa que o servidor DNS pensa que ele é o servidor DNS raiz da Internet, o que ele, certamente, não é. Apaque o . [ponto] para que qualquer coisa possa funcionar.

DNS dividido (Split DNS) e servidor espelhado

O propósito do DNS dividido (split DNS, também conhecido como horizonte dividido–***split horizon***) é apresentar uma visão diferente de seu domínio para os mundos interno e externo. Há mais de uma maneira de fazer um DNS dividido, mas, por razões de segurança, é recomendável que você tenha separados os conteúdos de seus servidores DNS interno e externo (cada um com base de dados distinta).

O split DNS pode permitir que clientes na rede de um campus resolvam endereços do domínio local do campus para os IPs do tipo RFC1918, enquanto o resto da Internet resolve os mesmos nomes para endereços IPs diferentes. Isto é conseguido tendo duas zonas em dois servidores DNS diferentes para o mesmo domínio.

Uma das zonas é usada pelos clientes da rede interna e a outra por usuários na Internet. Por exemplo, na rede abaixo, o usuário no campus de Makerere tem o endereço *http://www.makerere.ac.ug/* resolvido para 172.16.16.21, enquanto outro usuário qualquer na Internet tem o mesmo resolvido para 195.171.16.13.

O servidor DNS do campus, no caso acima, tem um arquivo de zona para makerere.ac.ug e está configurado como autoridade para este domínio. Em adição, ele serve como caching DNS para o campus de Makerere e todos os computadores no campus estão configurados para usá-lo como seu servidor DNS.

Os registros de DNS para o servidor do campus serão parecidos com estes:

```
makerere.ac.ug
www  CNAME webserver.makerere.ac.ug
ftp  CNAME ftpserver.makerere.ac.ug
mail CNAME exchange.makerere.ac.ug
mailserver A 172.16.16.21
webserver  A 172.16.16.21
ftpserver  A 172.16.16.21
```

Mas há outro servidor DNS na Internet que é realmente autoridade para o domínio *makerere.ac.ug*. Os registros DNS para a zona externa se parecerão com estes:

```
makerere.ac.ug
www  A 195.171.16.13
ftp  A 195.171.16.13
mail A 16.132.33.21
     MX mail.makerere.ac.ug
```

O DNS dividido não é dependente do uso de endereçamento RFC 1918. Um provedor de acesso à Internet africano poderia, por exemplo, hospedar os websites da universidade, mas também espelhar estes mesmos websites na Europa. Sempre que um cliente daquele provedor acessar o website, ele terá um endereço IP do provedor africano, mantendo o tráfego dentro do mesmo país. Quando visitantes de outros países acessarem o website, eles obtêm o endereço IP do site espelhado na Europa. Desta forma, visitantes internacionais não irão congestionar a conexão VSAT do provedor quando visitarem o site da universidade. Isto está se tornando uma solução atrativa, na medida em que a hospedagem de sites próximos ao backbone (links centrais, de alta velocidade) da Internet tornaram-se muito baratos.

Otimização do link de Internet

Como já foi mencionado, uma taxa de transmissão de 22 Mbps na rede pode ser atingida com o uso de equipamentos wireless no padrão 802.11g, sem necessidade de licenças. Esta largura de banda será ao menos uma ordem de magnitude acima daquela fornecida pela sua conexão com a Internet, e deve ser

suficientemente confortável para o suporte a muitos usuários simultâneos da Internet.

Mas se a sua conexão primária com a Internet for através de um link VSAT, você enfrentará alguns problemas de desempenho caso utilize apenas os parâmetros padrão do TCP/IP. Com a otimização do link VSAT você poderá melhorar significativamente os tempos de resposta no acesso a servidores na Internet.

Fatores TCP/IP em uma conexão via satélite

Freqüentemente, referimo-nos a um VSAT como uma rede de cano longo e grosso (***long fat pipe network***). Este termo tem a ver com os fatores que afetam o desempenho do TCP/IP em qualquer rede que tenha uma largura de banda relativamente grande, mas alta latência. A maioria das conexões à Internet na África e em outras partes do mundo em desenvolvimento são feitas com VSAT. Desta forma, mesmo que uma universidade tenha sua conexão à Internet fornecida por um provedor de acesso, a instruções fornecidas aqui também se aplicam, caso a conexão deste provedor à Internet for através de VSAT. A alta latência de redes via satélite é devida à longa distância até o satélite e à velocidade constante da luz. A distância adiciona cerca de 520 ms para o tempo de viagem de um pacote (***RTT*** – ***round-trip time***), comparado com um RTT típico entre a Europa e os Estados Unidos, de cerca de 140 ms.

Os fatores que mais significativamente impactam o desempenho do TCP/IP são ***RTT longo***, ***atrasos de entrega em largura de banda alta*** e ***erros de transmissão***.

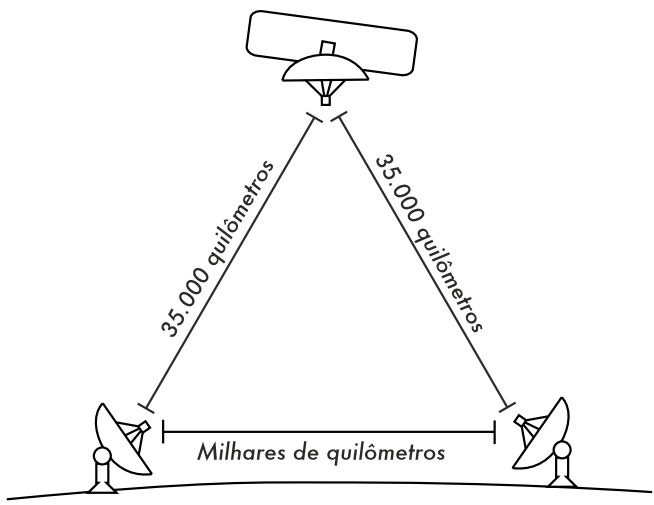


Figura 3.28: Devido à velocidade da luz e das longas distâncias envolvidas, um simples pacote de ping pode demorar mais de 520 ms para ser recebido em um link VSAT.

De maneira geral, sistemas operacionais que suportam implementações modernas de TCP/IP devem ser usados em redes via satélite. Estas implementações fornecem o suporte às extensões RFC 1323:

- A opção ***window scale*** para o suporte à largas janelas TCP (maiores que 64 KB);
- ***Selective acknowledgement*** (***SACK*** – reconhecimento seletivo) para permitir a rápida recuperação de erros de transmissão;
- ***Timestamps*** (registros de tempo) para o cálculo apropriado de RTT e valores de retransmissão por expiração de tempo (*timeout*) para o link em uso.

RTT Longo

Links via satélite tem um RTT médio de 520 ms para o primeiro enlace. O TCP usa um mecanismo de inicialização lenta (*slow start*) no estabelecimento da conexão a fim de descobrir os parâmetros apropriados para a mesma. O tempo gasto no estágio de inicialização lenta é proporcional ao RTT, o que significa que, em um link via satélite, este estágio dura mais tempo do que duraria em outros casos. Isto diminui drasticamente a taxa de transferência em conexões TCP de curta duração. Observa-se isto quando um website pequeno demora um tempo surpreendentemente longo para ser carregado, mas um arquivo grande é transferido a taxas de transferência aceitáveis.

Além disto, quando pacotes são perdidos o TCP entra em fase de controle de congestionamento (*congestion-control*) e, devido ao longo RTT, permanece nesta fase por um longo tempo, assim reduzindo a taxa de transferência tanto para conexões TCP de curta como de longa duração.

Produto do atraso de entrega em largura de banda alta (Large bandwidth-delay product)

A quantidade de dados em trânsito num link, em qualquer período de tempo, é produto da largura de banda e do RTT. Devido à alta latência do link via satélite, o atraso de produto em largura de banda é grande. O TCP/IP permite que o servidor remoto envie uma certa quantidade de dados sem a necessidade de confirmação de recebimento. Uma confirmação é necessária para qualquer dado recebido em uma conexão TCP/IP. Entretanto, o servidor remoto sempre pode enviar uma certa quantidade de dados sem que uma confirmação seja recebida, o que é importante para que se consiga uma boa taxa de transmissão em redes com conexões que apresentam grandes atrasos de entrega. Esta quantidade de dados é chamada de tamanho de janela TCP (***TCP window size***). Em implementações modernas de TCP/IP, o tamanho de janela é, usualmente, 64 KB.

Em redes via satélite, o valor do produto do atraso de entrega é importante. Para utilizar o link integralmente, o tamanho da janela de conexão deve ser igual ao do produto do atraso de entrega. Se o tamanho máximo de janela permitido é de 64 KB, a máxima taxa de transferência, teoricamente, atingível via satélite é (tamanho da janela) / RTT, ou 64 KB / 520 ms. Isto resulta em uma taxa de transferência máxima de 123 KB/s, ou seja, 984 kbps, independente do fato da capacidade do link ser muito maior.

Cada cabeçalho de segmento TCP contém um campo chamado ***advertised window*** (janela publicada), que especifica quantos bytes adicionais de dados o

receptor está preparado para aceitar. A *advertised window* é o tamanho disponível do *buffer* do receptor.

O remetente não tem a permissão de enviar mais bytes do que a advertised window. Para maximizar o desempenho, o remetente deve configurar o tamanho de seu buffer de envio e o receptor o tamanho de seu buffer de recepção para um número que não seja menor que o produto do atraso de entrega. Este tamanho de buffer tem o valor máximo de 64 KB na maioria das implementações TCP/IP modernas.

Para contornar este problema em pilhas TCP/IP em sistemas operacionais que não aumentam a janela além dos 64 KB, uma técnica conhecida como ***TCP acknowledgment spoofing*** ("trapaça" no reconhecimento de recepção) pode ser usada (veja Melhora de desempenho com PEP, abaixo).

Erros de transmissão

Em implementações mais antigas de TCP/IP, a perda de pacotes era sempre considerada como resultado de congestionamentos (ao invés de erros de conexão). Quando isto acontece, o TCP realiza manobras para evitar congestionamentos, passando a requerer três ACKs (reconhecimento de recepção) duplicados ou uma reinicialização lenta em caso de um timeout. Em função do longo RTT, uma vez que esta fase de controle de congestionamento é iniciada, links TCP/IP via satélite irão levar um longo tempo até que voltem ao nível de taxa de transmissão anterior ao problema. Desta forma, erros em um link via satélite têm um efeito muito mais sério na performance do que o TCP em links de latência baixa. Para contornar esta situação, mecanismos como o ***SACK*** (***Selective Acknowledgement***) foram desenvolvidos. O SACK especifica exatamente aqueles pacotes que foram recebidos, permitindo ao remetente retransmitir apenas os segmentos que foram perdidos por causa de erros de conexão.

O artigo sobre os detalhes de implementação do TCP/IP no Microsoft Windows 2000 declara:

> *"O Windows 2000 introduz o suporte para uma funcionalidade importante de desempenho, conhecida como SACK (Selective Acknowledgement), especialmente para conexões que usam grandes janelas TCP."*

O SACK vem sendo uma funcionalidade padrão nos kernels Linux e BSD por um longo tempo. Certifique-se de que seu roteador Internet e seu provedor tenham, ambos, suporte ao SACK.

Implicações para universidades

Se uma localidade tem uma conexão de 512 kbps para a Internet, a configuração padrão do TCP/IP provavelmente será suficiente, uma vez que uma janela de 64 KB é o bastante para uma velocidade de 984 kbps. Mas se a universidade tiver uma conexão com velocidade acima de 984 kbps, em alguns casos não será possível utilizar toda a largura de banda disponível em função das características de uma rede de cano longo e grosso (***long fat pipe network***) que discutimos acima. Estes fatores realmente evitam que uma única máquina ocupe integralmente a banda, o que não é ruim durante o dia, pois

muitas pessoas estão usando a rede. Mas se, por exemplo, há uma grande programação de downloads para a madrugada, o administrador da rede possivelmente irá querer que estes ocupem toda a largura de banda e, neste caso, as características da rede de cano longo e grosso podem ser um obstáculo. Isto também pode ser crítico se uma significativa parte de seu tráfego de rede é roteada por um único túnel, ou uma conexão VPN para a outra ponta do link VSAT.

Os administradores devem tomar as devidas precauções para que o uso de toda a banda possa ser conseguido através do ajuste fino do TCP/IP. Caso uma universidade tenha implementado uma rede onde todo o tráfego passe por um proxy (assegurado pela configuração da rede), então as únicas máquinas que farão a conexão com a Internet serão este proxy e os servidores de email.

Para mais informações, veja *http://www.psc.edu/networking/perf_tune.html*.

Melhora de desempenho com PEP (Performance-enhancing proxy)

A idéia de um proxy que melhore o desempenho da rede, ou ***PEP*** (***Performance-enhancing proxy***), é descrita no RFC 3135 (*http://www.ietf.org/rfc/rfc3135*) e consiste em um servidor proxy com um grande espaço em disco para o cache, usando as extensões RFC 1323 dentre outras funcionalidades. Um laptop tem uma sessão TCP com o PEP de um provedor de acesso. Este PEP comunica-se com o que está na outra extremidade da conexão via satélite, usando uma outra sessão TCP ou mesmo um protocolo proprietário. O PEP do provedor da conexão via satélite obtém os arquivos do servidor web. Desta forma, a sessão TCP é dividida e as características do link que afetam o desempenho do protocolo (os fatores da rede de cano grosso e longo) são contornados, usando ***TCP acknowledgment spoofing*** ("trapaça" no reconhecimento de recepção), por exemplo. Adicionalmente, o PEP faz uso de técnicas de busca antecipada de conteúdos (pre-fetching) para acelerar ainda mais o acesso à web.

Sistemas assim podem ser construídos do zero com o uso do Squid, por exemplo, ou adquiridos de uma variedade de fornecedores.

Mais informações

Enquanto a otimização de banda é um tema complexo e freqüentemente difícil, as técnicas descritas neste capítulo devem ajudar a reduzir as fontes mais claras de desperdício no uso da rede. Para utilizar da melhor forma possível a largura de banda disponível, você precisará definir uma boa política de acesso, configurar boas ferramentas para o monitoramento e análise, e implementar uma arquitetura que reforce os limites de utilização da rede.

Para mais informações sobre a otimização da largura de banda, consulte o livro *How to Accelerate Your Internet* (*http://bwmo.net/*), que está disponível livremente.