UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
ESCOLA POLITÉCNICA
DEPARTAMENTO DE ELETRÔNICA E DE COMPUTAÇÃO

Um Protótipo de Canal de TV Usando um Serviço de Difusão IP com Fonte Específica

Autor:

Igor Monteiro Moraes

Orientador:

Prof. Otto Carlos Muniz Bandeira Duarte, Dr.Ing.

Examinadores:

Célio Vinícius Neves de Albuquerque, Ph.D.

Luís Henrique Maciel Kosmalski Costa, Dr.

Prof. Marcelo Gonçalves Rubinstein, D.Sc.

Prof. Marcelo Luiz Drumond Lanza, M.Sc.

DEL
Março de 2004

À minha família.

Agradecimentos

Aos meus pais e demais familiares pelo carinho, compreensão e incentivo ao longo de toda a graduação.

Ao professor Otto, meu orientador, e ao recém-doutor Luís Henrique pelos conselhos durante a execução deste trabalho.

Aos amigos do GTA, em especial ao Doc, ao Pedro e ao Bernardo, por compartilharem seus computadores, permitindo a implementação da plataforma de testes.

Aos amigos da graduação e iniciação científica, Rafael e Marco, pela contribuição e incentivo para a realização deste projeto.

Aos professores do Departamento de Eletrônica e de Computação, principalmente aos professores Mauros, Lanza e Rezende, pela amizade e pelos conhecimentos adquiridos.

A todos que me incentivaram, contribuindo de forma direta ou indireta, para a minha formação profissional.

Resumo do Projeto Final apresentado ao DEL/UFRJ como parte dos requisitos necessários para a obtenção do grau de Engenheiro Eletrônico e de Computação.

Um Protótipo de Canal de TV Usando um Serviço de Difusão IP com Fonte Específica

Igor Monteiro Moraes

Março de 2004

Orientador: Otto Carlos Muniz Bandeira Duarte

Departamento: Engenharia Eletrônica e de Computação

As novas aplicações multimídias têm como característica comum a comunicação de grupo. Dentre estas novas aplicações, a distribuição de TV tem o potencial de ser a aplicação que vai disparar a implantação do serviço *multicast* na Internet, dado o enorme número de usuários envolvidos.

Neste trabalho são analisados dois dos serviços de comunicação de grupo propostos: o IP Multicast e o SSM (*Source-Specific Multicast*). Para demonstrar a viabilidade dos dois serviços como mecanismos de distribuição de TV, foi desenvolvida uma plataforma de testes, que simula um domínio típico da Internet e possui um protótipo de canal de TV.

Os testes realizados comprovam que o uso de um serviço *multicast* para a distribuição de TV na Internet proporciona uma redução na taxa de utilização da rede, se comparado com o uso de múltiplas conversações ponto-a-ponto para o mesmo propósito. Também foram realizados testes para avaliar o desempenho de alguns protocolos de roteamento *multicast*. Os resultados mostram que o serviço SSM elimina o problema da alocação global de endereços, facilita o controle de acesso ao grupo e garante a integridade das aplicações de distribuição de vídeo, requisitos necessários para a TV na Internet.

Palavras Chave

Redes de Computadores

Internet

Distribuição de TV

IP Multicast

SSM

Lista de Acrônimos

ADSL :	*Asymmetric Digital Subscriber Line*;
BGP :	*Border Gateway Protocol*;
BGMP :	*Border Gateway Multicast Protocol*;
CC :	*Contributed Source Counter*;
CSRC :	*Contributed Source*;
DR :	*Designated Router*;
DVMRP :	*Distance Vector Multicast Routing Protocol*;
EXPRESS :	*Explicitly Requested Single Source Multicast*;
FTP :	*File Transfer Protocol*;
HTTP :	*Hyper Text Transfer Protocol*;
IANA :	*Internet Assigned Numbers Authority*;
IETF :	*Internet Engineering Task Force*;
IGMP :	*Internet Group Management Protocol*;
IGMPv1 :	*Internet Group Management Protocol*, versão 1;
IGMPv2 :	*Internet Group Management Protocol*, versão 2;
IGMPv3 :	*Internet Group Management Protocol*, versão 3;
IP :	*Internet Protocol*;
IPv4 :	*Internet Protocol*, versão 4;
MAAA :	*Multicast Address Allocation Architecture*;
MBone :	*Multicast Backbone*;
MBGP :	*Multiprotocol Extensions for BGP-4*;
MPEG :	*Moving Picture Experts Group*;

MSDP :	*Multicast Source Discovery Protocol*;
NTP :	*Network Time Protocol*;
NTSC :	*National Television System Committee*;
OSPF :	*Open Shortest Path First*;
PIM :	*Protocol Independent Multicast*;
PIM-SM :	*Protocol Independent Multicast – Sparse Mode*;
PIM-DM :	*Protocol Independent Multicast – Dense Mode*;
PIM-SSM :	*Protocol Independent Multicast – Source-Specific Multicast*;
QoS :	*Quality of Service*;
RIP :	*Routing Information Protocol*;
RP :	*Rendez-vous Point*;
RPF :	*Reverse Path Forwarding*;
RR :	*Receiver Report*;
RTP :	*Real Time Protocol*;
RTCP :	*Real Time Control Protocol*;
RTSP :	*Real Time Streaming Protocol*;
SAP :	*Session Announcement Protocol*;
SDP :	*Session Description Protocol*;
SDR :	*Session Directory Tool*;
SMTP :	*Simple Mail Transfer Protocol*;
SR :	*Sender Report*;
SSM :	*Source-Specific Multicast*;
SSRC :	*Synchronization Source*;
TBT :	*Truncated Broadcast Trees*;
TCP :	*Transmission Control Protocol*;
TTL :	*Time-To-Live*;
UDP :	*User Datagram Protocol*;
URL :	*Uniform Resource Locator*;
VIC :	*Videoconference Tool*;
VLC :	*VideoLAN Client*;

VLS : *VideoLAN Server*;

VLAN : *Virtual Local Area Network*;

XML : *Extensible Markup Language*;

WB : *Whiteboard Tool.*

Sumário

Lista de Figuras

Lista de Tabelas

Capítulo I

Introdução

Hoje em dia, a maior parte das aplicações da Internet constituem conversações ponto-a-ponto. Entretanto, as novas aplicações multimídias, como a videoconferência, o ensino à distância e a distribuição de televisão, são caracterizadas pelo envio de uma mesma informação para múltiplos receptores. No caso da distribuição de TV, o número de receptores pode chegar aos milhões, o que torna a utilização de múltiplos canais ponto-a-ponto inviável em termos dos recursos da rede e da capacidade de processamento das estações [1]. Por isso, é necessário um serviço de distribuição *multicast* eficiente.

Proposto há mais de dez anos, o IP Multicast [2] vem sendo implantado lentamente como serviço nativo na Internet, em virtude da complexidade para configurar e gerenciar o conjunto de protocolos de rede necessários à sua implementação. A razão desta complexidade deve-se ao fato do modelo de serviço IP Multicast definir um *grupo* como uma conversação, onde existem N fontes e M receptores, da qual qualquer estação pode participar sem exigência de autorização. O modelo também define que uma estação pode pertencer a vários grupos, podendo entrar ou sair de um grupo a qualquer instante, e que uma fonte pode enviar dados, pertencendo ou não ao grupo.

Um grupo é identificado por um endereço IP pertencente a uma faixa reservada, chamado de endereço IP *multicast*. O endereço IP *multicast* corresponde ao endereço de um grupo de estações e não ao endereço de uma máquina específica. Ele é alocado dinamicamente e permanece válido durante toda a sessão *multicast*.

Devido à complexidade e, conseqüentemente, à lenta implantação do IP Multicast na Internet, novas alternativas vêm sendo propostas para simplificar o serviço. Uma das alternativas que vem recebendo um grande destaque é o serviço SSM (*Source-Specific Multicast*) [3], um serviço de difusão IP com fonte específica.

Atualmente em fase final de padronização pelo IETF (*Internet Engineering Task Force*), o serviço SSM foi inspirado no protocolo EXPRESS (*EXPlicitly REquested Single Source multicast*) [4]. O SSM usa o conceito de canal proposto neste protocolo. Um canal possui uma única fonte por grupo e é identificado por um par de endereços (fonte,grupo). Esta definição reduz a conversação *multicast* para um-para-muitos, $1XM$, simplificando a arquitetura do serviço.

A proposta de um serviço *multicast* de fonte específica se justifica, pois a maioria das aplicações de grupo conhecidas atualmente possui uma única fonte. Além disso, nas aplicações com múltiplas fontes, o número de fontes é pequeno, caso da videoconferência, o que torna viável o uso de múltiplos canais de fonte específica. Os exemplos de aplicações de fonte única de maior destaque são a distribuição de rádio e televisão.

Dado o enorme número de usuários envolvidos, a distribuição de TV na Internet será a aplicação que vai acelerar a implantação de um serviço *multicast* em larga escala. Por ser uma aplicação de fonte única, é natural o uso do serviço SSM para o transporte eficiente do tráfego de televisão na Internet.

I.1 Objetivos

O principal objetivo deste trabalho é demonstrar a viabilidade do serviço SSM como mecanismo de distribuição de TV na Internet. Para isso, foi desenvolvida uma plataforma de testes que simula um domínio típico da Internet, composta por roteadores e estações com suporte aos protocolos necessários para implementar o serviço SSM.

A plataforma conta ainda com um protótipo de canal de TV na Internet. O protótipo é composto por um único computador, equipado com uma placa de captura e com uma câmera de vídeo. Este computador desempenha as funções de servidor de vídeo e de máquina de codificação. O servidor é capaz de distribuir vídeo nos formatos MPEG-2, usado por todos os padrões de transmissão de TV digital terrestre e via satélite, e MPEG-4, desenvolvido para a transmissão de TV interativa na Internet. Todas as estações possuem programas clientes de vídeo compatíveis com os padrões MPEG-2 e MPEG-4. Durante os testes, também foi usada uma aplicação de videoconferência, o VIC (*Videoconference Tool*).

A análise da viabilidade do serviço SSM como mecanismo de distribuição de TV foi realizada através de testes práticos, sendo dividida em três partes. Na primeira parte foi feita uma comparação entre o uso do serviço *multicast* e o uso de múltiplas conversações ponto-a-ponto para a distribuição de TV. A métrica usada nesta comparação foi a taxa de utilização da rede. Na segunda, foram analisados os mecanismos da alocação de endereços e de controle de acesso, proporcionados pelo IP Multicast e pelo SSM. Na terceira parte, comparou-se o desempenho de alguns protocolos de roteamento *multicast*. Foram analisados os protocolos DVMRP (*Distance Vector Multicast Routing Protocol*), PIM-DM (*Protocol Independent Multicast - Dense Mode*) e PIM-SM (*Protocol Independent Multicast - Sparse Mode*), usados pelo serviço IP Multicast, e o protocolo PIM-SSM (*Protocol Independent Multicast - Source-Specific Multicast*), usado no serviço SSM.

Os protocolos de roteamento são os responsáveis pela criação da árvore de distribuição do tráfego *multicast*. Cada um dos protocolos constrói um tipo de árvore usando diferentes mecanismos, enviando pacotes de dados e de controle em diferentes quantidades e em diferentes partes da rede. Por isso, a avaliação do desempenho dos protocolos de roteamento *multicast* é fundamental para determinar qual é o protocolo mais adequado à aplicação de distribuição de TV. Os testes realizados simularam uma transmissão de um programa de TV na Internet. Enquanto a fonte distribuía o vídeo e os clientes entravam e deixavam o grupo, os enlaces da rede eram monitorados. As métricas usadas para comparar o desempenho dos protocolos foram a complexidade de configuração do protocolo, a taxa de utilização da rede durante a construção da árvore de distribuição, o número de pacotes de controle de roteamento e a capacidade de reconfiguração da árvore de distribuição.

I.2 Organização do Texto

Este trabalho está organizado da seguinte forma. No Capítulo II são apresentadas as principais características da distribuição de vídeo em redes de comutação por pacotes. Nos Capítulo III e IV, respectivamente, os serviços IP Multicast e SSM são discutidos. No Capítulo V são apresentados, em detalhes, o protótipo de canal de TV e a plataforma de testes desenvolvida. Ainda nesse capítulo são descritas as ferramentas usadas no desenvolvimento deste trabalho. Detalhes referentes aos testes e à análise dos resultados são abordados no Capítulo VI. Por fim, no Capítulo VII serão apresentadas as conclusões sobre este trabalho e as considerações sobre trabalho futuros.

Capítulo II

Distribuição de TV na Internet

A Internet é uma rede baseada na técnica de comutação de pacotes para a transmissão de dados. Ao contrário da técnica de comutação de circuitos, na comutação de pacotes o meio físico é compartilhado pelos usuários, o que permite o uso mais eficiente dos recursos da rede. Para a transmissão de dados, a comutação de pacotes é ideal, porém para a transmissão de mídias com restrições de tempo e perda de pacotes, como é o caso da distribuição de vídeo digital codificado, ela apresenta alguns problemas.

Este capítulo descreve a atual arquitetura da Internet e aborda as principais características da distribuição de vídeo em redes de comutação de pacotes. Também são apresentados os novos protocolos para o transporte do tráfego multimídia e os modelos de negócio para a distribuição de TV na Internet.

II.1 A Arquitetura da Internet

No início da década de 1990, a Internet era uma pequena rede de pesquisa usada por poucos usuários. Atualmente, a Internet consiste de uma rede de alcance mundial que interconecta diversas redes heterogêneas, onde trafegam aplicações com diferentes requisitos. Isto implica que arquitetura da Internet deve ser flexível, para suportar as mais diversas aplicações, e independente do tipo de tecnologia física empregada pelas redes que ela interconecta.

A Internet é uma rede de comutação de pacotes. Neste tipo de rede, a informação é dividida em pacotes. O protocolo de rede usado não faz estabelecimento prévio de conexão, a comunicação de um usuário pode interferir na dos outros e os pacotes podem chegar desordenados ao destinatário.

Como o meio físico é compartilhado, é necessária a adição de um cabeçalho em cada pacote. Isto aumenta a complexidade dos nós intermediários da rede, que precisam analisar o conteúdo do cabeçalho de cada pacote para encaminhá-lo corretamente.

Além do atraso de propagação no meio, tem-se os atrasos de encapsulamento, de desencapsulamento e de processamento dos pacotes e o atraso de espera dos pacotes nas filas dos roteadores. Como não há conexão e a banda passante para cada usuário é alocada dinamicamente, de acordo com a disponibilidade da rede, não há garantia de vazão, atraso ou variação do atraso.

O modelo de referência TCP/IP, no qual é baseada a arquitetura da Internet, é composto por quatro camadas [5], como mostra a Figura II.1. O modelo recebeu este nome em virtude dos dois principais protocolos da Internet: o TCP (*Transmission Control Protocol*) e o IP (*Internet Protocol*).

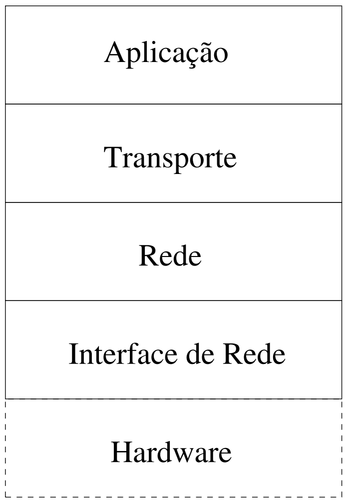


Figura II.1: A arquitetura TCP/IP.

A camada de aplicação representa o usuário final na arquitetura TCP/IP. Ela contém uma série de protocolos largamente usados pelos usuários. O FTP (*File Transfer Protocol*), usado para a transferência de arquivos, e o SMTP (*Simple Mail Transfer Protocol*), usado para a troca de mensagens de correio eletrônico, são exemplos de protocolos da camada de aplicação.

A camada de transporte provê um serviço fim-a-fim, que permite a comunicação entre entidades pares localizadas nas estações fonte e destino. O TCP e o UDP (*User Datagram Protocol*) são protocolos da camada de transporte.

O TCP é um protocolo orientado a conexão e confiável, pois garante que a mensagem originada no emissor chegará ao seu destino sem erros. Para garantir a entrega da mensagem sem erros, cada segmento possui um número de seqüência. Além disso, o receptor deve enviar reconhecimentos ao emissor, informando-o sobre a boa recepção da informação. Eventuais perdas de informação são retransmitidas. O TCP também fornece um mecanismo de controle de fluxo, através do uso de janelas deslizantes para transmissão de segmentos em avanço ao reconhecimento.

Já o UDP é um protocolo sem conexão e não confiável, pois não possui nenhuma garantia de entrega da mensagem enviada. O protocolo UDP não usa números de seqüência, nem exige o envio de reconhecimentos por parte do receptor ao emissor. Cada mensagem é tratada independentemente das outras e pode chegar ao receptor fora de ordem. Entretanto, como o UDP não exige conexão, nem controle de erros, nem controle de fluxo, o processamento requerido é bem menor que o do TCP. Por isso, o UDP é o protocolo de transporte mais usado pelas aplicações multimídias.

A camada de rede é a responsável por determinar o melhor caminho para o envio dos pacotes, por encaminhar os pacotes até o destino e por interconectar redes das mais diversas tecnologias. O serviço de seleção do melhor caminho, chamado de roteamento, é responsabilidade dos protocolos de roteamento, como o RIP e o OSPF. O encaminhamento dos pacotes e a interconexão de redes são funções do IP.

Os nós responsáveis por interconectar redes distintas são chamados de roteadores. Ao chegar a um roteador, cada pacote tem o endereço de destino, contido no seu cabeçalho, analisado. Se o endereço de destino corresponder ao endereço de uma das interfaces do próprio roteador, ele é enviado à camada de transporte. Do contrário, é escolhida uma interface de saída por onde o pacote é encaminhado até o próximo roteador. Este processo repete-se, a cada roteador do caminho, até que o pacote chegue ao seu destino.

O protocolo IP é responsável pelo encaminhamento dos pacotes. O IP oferece um serviço não confiável, sem estabelecimento prévio de conexão, em que os pacotes são encaminhados de acordo com o modelo de melhor esforço [5].

É dito que o serviço oferecido pelo IP é de melhor esforço, porque todos os roteadores que constituem a rede se esforçam ao máximo para que todos os pacotes alcancem seu destino final. Ao chegarem a um roteador, os pacotes são tratados sem distinção, da melhor maneira possível, respeitando a ordem de chegada.

A camada interface de rede é responsável por transmitir sobre o meio físico os datagramas IP provenientes da camada de rede. Tanto o *driver* quanto o conjunto de protocolos desta camada dependem da tecnologia de rede empregada.

II.2 Distribuição de Vídeo em Redes

A distribuição de vídeo traz novos desafios para as atuais redes de comutação de pacotes. Com o aumento da capacidade de processamento dos computadores pessoais e o surgimento de novas aplicações multimídias, a Internet tende a ser cada vez mais usada para transportar o tráfego multimídia. Por não haver reserva de recursos para cada usuário da rede, nem diferenciação no tratamento dos pacotes encaminhados pelos roteadores, a qualidade de serviço (QoS - *Quality of Service*) exigida por estas novas aplicações não é atendida pela atual arquitetura da Internet. Não há garantia de vazão, atraso ou variação do atraso (*jitter*).

O vídeo é uma mídia contínua que requer uma grande largura de banda passante e é sensível ao atraso e à variação do atraso. Apesar de ser uma mídia que apresenta alta redundância espacial e temporal, o vídeo também é sensível à taxa de perda de pacotes. Quando codificado, as informações redundantes contidas no vídeo são reduzidas para diminuir o volume de informação e, conseqüentemente, o espaço ocupado em memória e o tempo de transmissão. No caso da codificação MPEG-4, uma taxa de perda de 2,0% dos pacotes já degrada totalmente a qualidade da transmissão do vídeo [6].

II.2.1 O Problema da Heterogeneidade dos Receptores

A distribuição de vídeo na Internet difere significativamente da distribuição convencional de televisão. Na distribuição convencional de TV, o sinal transmitido tem banda passante fixa e os receptores são homogêneos. A qualidade é fixa e a boa recepção só depende do nível de sinal que chega aos receptores. Já na distribuição de vídeo pela Internet, a informação de vídeo é transmitida em pacotes, a banda passante da rede é compartilhada dinamicamente pelas aplicações em andamento e os receptores são heterogêneos [7]. A Figura II.2 ilustra o problema da heterogeneidade dos receptores na Internet.

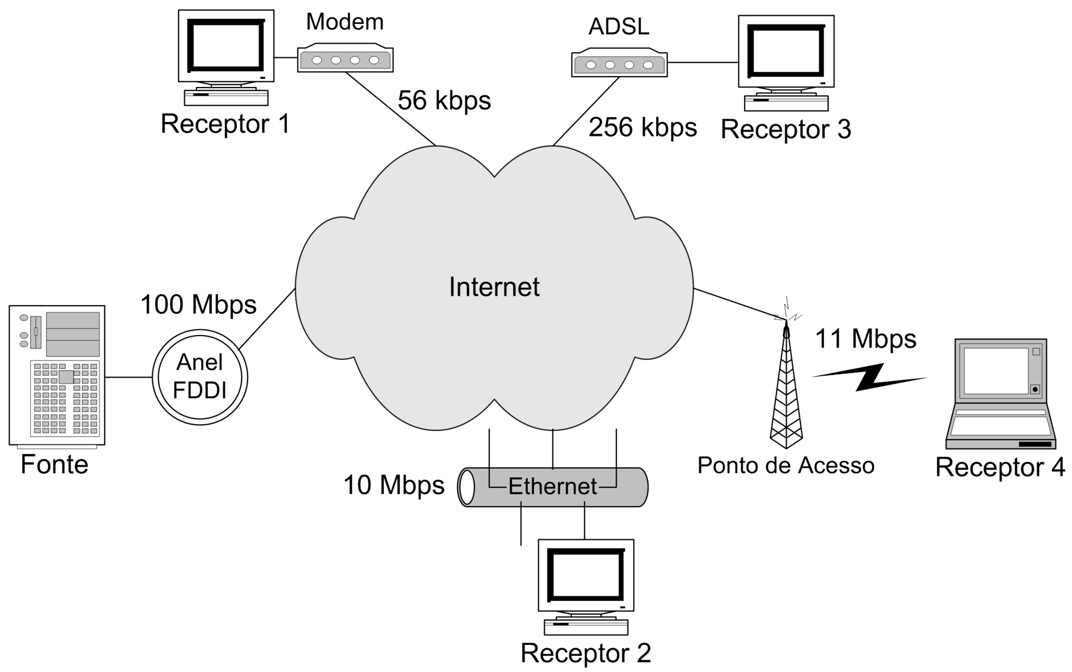


Figura II.2: A heterogeneidade dos receptores.

A Internet é formada por redes com características distintas de banda passante. Os usuários podem estar conectados via modem, *cable modem*, ADSL (*Asymmetric Digital Subscriber Line*), redes sem fio ou rede local. Além disso, deve-se considerar que a banda passante da rede é compartilhada e que os computadores de cada usuário possuem capacidades de processamento diferentes. Essa diversidade na capacidade dos receptores impede que uma única taxa de transmissão atenda satisfatoriamente a todos os receptores da Internet.

Uma solução do tipo "melhor denominador comum" não é suficiente. Se a taxa de transmissão for adaptada ao receptor de maior capacidade, haverá congestionamento nos enlaces receptores de menor capacidade. Já se a taxa de transmissão for adaptada ao receptor de menor capacidade, os receptores de maior capacidade receberão um vídeo, cuja qualidade será menor do que a que eles podem receber. Portanto, é necessário que receptores com capacidades diferentes recebam vídeos com níveis de qualidade diferentes.

II.2.2 Propostas para Distribuição de Vídeo na Internet

Existem várias propostas para distribuir vídeo na Internet com diferentes níveis de qualidade. A primeira delas é conhecida como *simulcast* [8]. Nesta técnica, a fonte transmite vários fluxos independentes do mesmo vídeo com qualidades distintas, como mostra a Figura II.3. De acordo com suas restrições de banda passante, cada receptor escolhe qual dos fluxos deseja receber. Apesar de proporcionar a distribuição de vídeo com diferentes níveis de qualidade, na técnica *simulcast* a banda passante da rede é usada de forma ineficiente, já que são transmitidos vários fluxos do mesmo vídeo.

Outra técnica proposta é a transcodificação [9]. Nesta técnica, a fonte transmite a uma taxa elevada um único fluxo de vídeo codificado. Cabe aos nós intermediários transcodificar - decodificar e recodificar - o fluxo de vídeo nos pontos de gargalo da rede. O custo computacional exigido pela técnica de transcodificação é elevado. Além disso, há desperdício de banda passante entre a fonte e os pontos de gargalo da rede.

Uma outra proposta para a distribuição de vídeo com diferentes níveis de qualidade

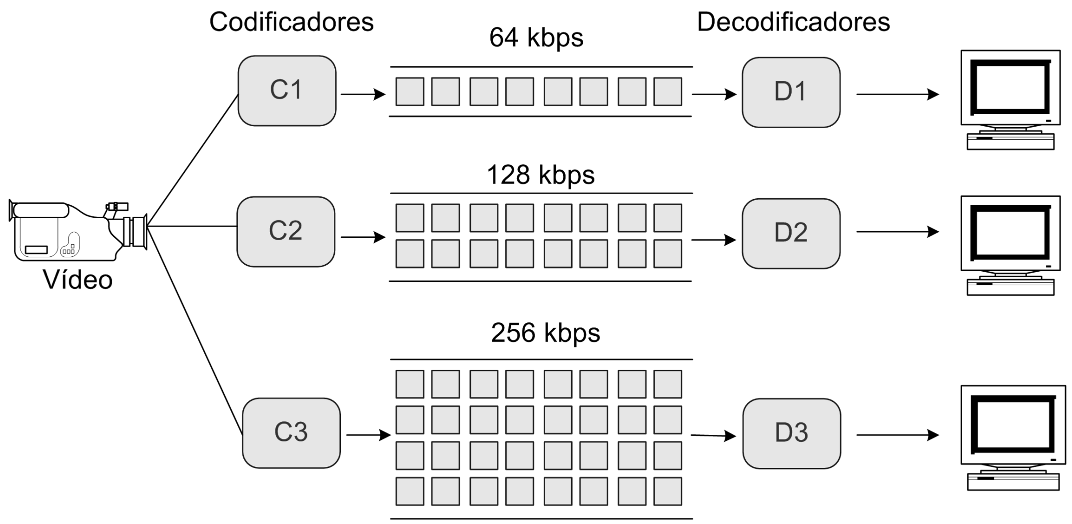


Figura II.3: A técnica *simulcast* para distribuição de vídeo.

é a técnica de transmissão de vídeo em camadas [10, 11]. Nesta técnica, um codificador comprime o vídeo em uma ou mais camadas hierárquicas com diferentes prioridades. A camada que contém as informações essenciais do fluxo de vídeo é chamada de camada base. Esta é a camada de mais alta prioridade. Todo receptor deve ser capaz de receber pelo menos esta camada. Camadas adicionais com prioridade decrescente podem ser codificadas para refinar a qualidade da camada base. Essas camadas de refinamento são somadas à camada base, ou a outras camadas de refinamento com maior prioridade, aumentando progressivamente a qualidade da informação de vídeo reconstruída. O refinamento pode estar relacionado à resolução, à taxa de quadros ou à relação sinal-ruído com que o fluxo de vídeo chega ao receptor. A Figura II.4 ilustra o processo de codificação em camadas.

A cada uma das camadas é associado um endereço *multicast*. Dessa forma, cada receptor pode-se associar a um ou mais grupos *multicast* de acordo com as suas limitações de banda passante. Cada uma das camadas é transmitida sobre uma árvore de distribuição diferente, o que requer o uso de um serviço de distribuição *multicast* eficiente.

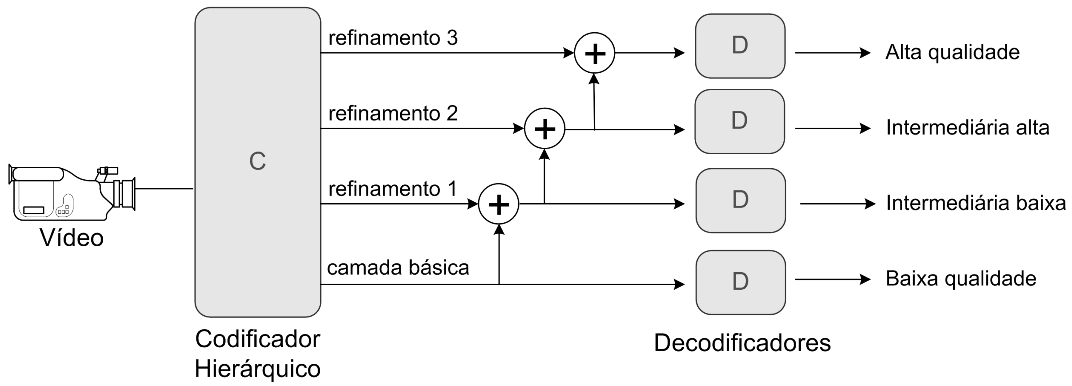


Figura II.4: A codificação em camadas.

II.3 Protocolos de Tempo Real para Aplicações Multimídias

O UDP é o protocolo de transporte mais usado pelas aplicações multimídias, uma vez que não exige o envio de reconhecimentos do receptor para o emissor e por não efetuar controle de fluxo e de erros. No entanto, mídias contínuas têm restrição de tempo e requerem ordenação da informação. Para fornecer esses serviços foram desenvolvidos os protocolos de transporte e controle para aplicações multimídias de tempo real.

II.3.1 O Protocolo RTP

O protocolo RTP(*Real Time Protocol*) fornece um serviço de entrega fim-a-fim para dados com restrições de tempo, como áudio e vídeo [12]. O RTP fornece mecanismos para a ordenação dos pacotes e estampilha de tempo para controle de cadência ou descarte de dados "fora do prazo" de entrega de pacotes. Ele não reserva recursos, não garante qualidade de serviço e não garante a entrega de pacotes. Na maioria das aplicações, o RTP é transportado encapsulado pelo UDP. Isto ocorre, pois o RTP usa os serviços de multiplexação e verificação de erros oferecidos pelo UDP. A Figura II.5 mostra o formato do cabeçalho do RTP.

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

V=2	P	X	CC	M	tipo de dados	número de seqüência
estampilha de tempo						
identificador de fonte de sincronização (SSRC)						
identificadores de fontes contribuintes (CSRC) ...						

Figura II.5: O cabeçalho do protocolo RTP.

Os primeiros 20 bytes do cabeçalho estão presentes em todos os pacotes RTP. Os dois primeiros bits identificam a versão do protocolo. O bit seguinte, quando setado em um, indica que no fim o pacote existem um ou mais bytes de enchimento que não fazem parte dos dados transportados e devem ser ignorados. O bit X indica que o cabeçalho possui uma extensão. O campo CC contém o número de identificadores de fontes contribuintes (CSRC - *Contributed Source*) que estão após a parte fixa do cabeçalho. Quando um pacote RTP transporta informações produzidas por mais de uma fonte, cada uma destas fontes é chamada de contribuinte. O bit M é interpretado de acordo com o perfil usado pela aplicação. O campo *tipo de dados* identifica o formato dos dados transportados pelo pacote RTP e determina como este pacote deve ser tratado pela aplicação. Os próximos 16 bits representam o *número de seqüência* de cada pacote RTP. O número de seqüência presente no cabeçalho permite que o receptor reconstrua a seqüência de pacotes enviada pelo emissor. O próximo campo contém a *estampilha de tempo*, que é usada para sincronizar os participantes de uma sessão RTP e calcular o atraso e a variação do atraso de cada pacote. Uma sessão RTP é definida como a associação entre um conjunto de participantes que comunicam-se através do protocolo RTP. Os 32 bits seguintes indicam a fonte de sincronização (SSRC -*Synchronization Source*), que é a fonte responsável pela transmissão do pacote RTP. Este identificador independe do endereço de rede. Caso o campo CC não seja nulo, ao final da parte fixa do cabeçalho estão listadas as fontes contribuintes.

II.3.2 O Protocolo RTCP

O protocolo RTCP (*Real Time Control Protocol*) monitora a qualidade de serviço da transmissão e coleta informações sobre os participantes de uma sessão RTP. Para tanto, pacotes de controle são transmitidos, periodicamente, a todos os participantes de uma sessão [12].

Cada pacote RTCP começa com uma parte fixa, similar à dos pacotes RTP. Os dois principais pacotes de controle do RTCP são os relatórios do emissor (SR - *Sender Report*) e do receptor (RR -*Receiver Report*).

O relatório do emissor é enviado aos receptores e informa a quantidade total de dados transmitidos e contém uma estampilha de tempo para sincronização. Já o relatório do receptor contém informações instantâneas e cumulativas sobre a variação do atraso e a taxa de perda de cada bloco de dados enviado por uma determinada fonte. As informações sobre a qualidade da transmissão, enviadas ao emissor, são usadas por mecanismos externos ao protocolo para controle de fluxo e congestionamento. Na maioria dos casos, os mecanismos são implementados pela aplicação.

II.3.3 O Protocolo RTSP

O RTSP (*Real Time Streaming Protocol*) é um protocolo da camada de aplicação que atua como um controle remoto de rede para servidores multimídias [13]. Ele permite a reprodução, a pausa, a interrupção, o avanço e o retrocesso rápidos de fluxos multimídias.

Os fluxos controlados pelo RTSP devem usar o protocolo RTP, mas a operação do RTSP não depende do mecanismo de transporte usado na distribuição da informação. A sintaxe e a operação do RTSP são, propositalmente, similares à versão 1.1 do protocolo HTTP (*Hyper Text Transfer Protocol*), usado para visualizar páginas na Internet. No entanto, um servidor RTSP precisa manter estados armazenados, o que não ocorre no HTTP, e tanto um servidor quanto um cliente RTSP podem enviar requisições.

A requisição do conteúdo multimídia ao servidor é feita através de uma URL (*Uniform Resource Locator*) RTSP, cuja sintaxe é similar à do protocolo HTTP. Um exemplo de uma URL RTSP seria: `rtsp://videoserver/video1.mp4`.

II.4 Modelos de Negócio e Produção de Conteúdo para a TV na Internet

Um modelo de negócio define a estratégia de operação de uma empresa para alcançar suas metas e objetivos, isto é, a acumulação de valor e a satisfação dos clientes [14].

Desde o início da década, vários experimentos de distribuição de conteúdo de entretenimento vêm sendo realizados na Internet. A maior parte destes experimentos envolveu conteúdo produzido exclusivamente para a Internet, como curtas-metragens e pequenas séries. Outros experimentos foram realizados com filmes já exibidos no cinema e alguns programas de TV gravados. Uma grande diversidade de modelos de negócio foi observada: propaganda, patrocínio, promoção de produtos e eventos, compras *on-line*, eventos em *pay-per-view*, entre outros. Apesar do número de experimentos realizados, atualmente, não é possível determinar qual, ou quais, destes modelos de negócio será adotado para a distribuição de TV na Internet.

Se comparada à distribuição tradicional de televisão no aspecto econômico, a TV na Internet apresenta uma série de vantagens [15]:

- baixo custo de distribuição e ampla área de cobertura, já que a Internet é uma rede de interconexão global e não há necessidade da construção de uma nova infra-estrutura;
- maior interatividade, pois a Internet possibilita aos seus usuários o recebimento de informações de acordo com os seus próprios interesses;
- maior eficiência em campanhas publicitárias e de patrocínio, que poderão ser direcionadas a um segmento específico de usuários;

- maior eficiência na venda direta ao consumidor, já que custos poderão ser reduzidos com a eliminação de intermediários;
- baixo custo de reprodução e gravação do conteúdo, uma vez que estas duas ações serão realizadas no próprio computador, sem a necessidade da compra de um novo equipamento.

Quanto a produção de conteúdo para a TV na Internet, existem duas vertentes. Na primeira, o conteúdo seria produzido exclusivamente para a distribuição na Internet. Na segunda, o conteúdo das grandes redes de TV aberta e das TVs por assinatura seria retransmitido na Internet, inclusive os eventos ao vivo. Nos dois casos, a tendência é que os distribuidores cada vez mais se especializem em um tipo de conteúdo específico, o que já ocorre nas TVs por assinatura a cabo e via satélite.

Como será visto adiante, o uso de um serviço de distribuição *multicast* para a TV na Internet impõem dificuldades de tarifação. Além disso, como os receptores são heterogêneos, eles receberão qualidades diferentes de vídeo e, portanto, deverão pagar de forma diferenciada.

Capítulo III

O Serviço IP Multicast

Proposto há mais de dez anos, o serviço IP Multicast se baseia em um modelo de comunicação no qual um conjunto de estações pode ser agregado em um *grupo*. Um grupo é um conjunto de zero ou mais estações identificadas por um único endereço IP de destino [2]. Portanto, ao enviar um datagrama para o endereço IP do grupo, uma estação estará enviando datagramas para todos os membros desse grupo.

Neste capítulo será apresentado o modelo de serviço IP Multicast, assim como será descrito o funcionamento dos principais protocolos de roteamento *multicast* e do protocolo de gerenciamento de grupo, o IGMP.

III.1 O Modelo de Serviço IP Multicast

No modelo de serviço proposto pelo IP Multicast, o grupo é uma conversação aberta, em que qualquer estação pode participar sem exigência de autorização. Os membros de um grupo são dinâmicos, uma estação pode entrar e sair de um grupo a qualquer instante. Não há restrições quanto à localização ou ao número de estações em um grupo. Uma estação pode pertencer a mais de um grupo simultaneamente e não precisa ser um membro do grupo para enviar datagramas para esse grupo. A identidade dos membros de um grupo também não é conhecida nem pela fonte e nem pelos receptores.

Existem dois tipos de grupos: permanentes ou transitórios. Um grupo permanente possui um endereço IP reservado associado administrativamente. Vale ressaltar que o endereço do grupo é permanente, não os seus membros. Um grupo transitório é criado quando uma fonte de tráfego escolhe um endereço *multicast* e envia dados para este endereço. Mais detalhes sobre a alocação de endereços serão vistos na Seção III.2. Para que um grupo transitório seja destruído, basta que a fonte pare de enviar dados para o endereço *multicast* associado ao grupo.

Além de aberta, a conversação em um grupo *multicast* pode ser: um-para-muitos, muitos-para-muitos e muitos-para-um. A distribuição de TV na Internet é um exemplo de conversação um-para-muitos, conhecida também como $1 \times N$, em que uma fonte envia dados para N receptores. Os jogos distribuídos são exemplos de conversações em que M fontes enviam dados para N receptores, chamadas de muitos-para-muitos, ou $M \times N$. Já os exemplos de conversação muitos-para-um são as votações e leilões *on-line*. Neste tipo de conversação, também chamada de $M \times 1$, M fontes enviam dados para apenas um receptor.

Uma estação com suporte ao IP Multicast pode entrar e sair de um grupo, assim como pode enviar datagramas IP a qualquer grupo de estações. Para que uma estação possa usar o serviço IP Multicast, ela deve implementar o protocolo IGMP (*Internet Group Management Protocol*) e as extensões da interface de serviço IP e da interface de rede local especificadas em [2].

A aplicação é responsável por sinalizar à camada de rede a intenção de uma estação se conectar a um grupo. A camada de rede, então, verifica se a estação já é integrante ou não do grupo solicitado. Caso a estação ainda não seja um membro do grupo, é enviada uma mensagem IGMP relatório (ver Seção III.3.2) à rede local.

Quando a uma estação se conecta a um grupo, a camada de rede e as camadas inferiores são reprogramadas para permitir a recepção de pacotes *multicast* endereçados a este grupo. Além disso, o endereço IP Multicast de um grupo é mapeado em um endereço de nível mais baixo.

Uma estação se conecta a um grupo *multicast* em uma interface, portanto, uma estação com várias interfaces pode escolher a interface através da qual se conectará ao grupo.

É importante ressaltar que um datagrama *multicast* é entregue a todos os membros do seu grupo de estações de destino com a mesma confiabilidade de um datagrama *unicast*, segundo o modelo de serviço de melhor esforço oferecido pelo protocolo IP. Dessa forma, não há garantia de que o datagrama chegará sem erros aos membros do grupo de destino ou na mesma ordem em que foram transmitidos.

III.2 Endereçamento

Um endereço *multicast* é dinamicamente escolhido por uma fonte de dados e válido durante toda a sessão *multicast*.

Na versão 4 do protocolo IP, um grupo *multicast* é identificado por um endereço IP classe D, ou seja, endereços cujos 4 bits mais significativos valem "`1110`". Isto corresponde à faixa de endereços que vai de `224.0.0.0` a `239.255.255.255`.

Os endereços da faixa de `224.0.0.1` a `224.0.0.255` foram reservados pelo IANA (*Internet Assigned Numbers Authority*) para serviços específicos. Estes endereços são de escopo local ao enlace e não devem ser encaminhados por roteadores, independente do valor do campo TTL (*Time-To-Live*) [1]. A faixa de endereços dinâmicos vai de `224.0.1.0` a `239.255.255.255`.

O endereço `224.0.0.0` não pode ser atribuído a nenhum grupo. Além disso, não existe um endereço IP *multicast* atribuído a todas as estações da Internet.

Como um endereço *multicast* não é atribuído por nenhuma entidade reguladora, ele é dinamicamente escolhido pela fonte, existe o risco de que várias fontes, que não desejam participar do mesmo grupo, escolham o mesmo endereço para transmitir seus dados. Sendo assim, dependendo do escopo de cada tráfego, as aplicações poderão ser corrompidas.

Para que não haja interferência entre as aplicações, é necessário um mecanismo de alocação de endereços que evite as colisões. Entretanto, visto que não há uma estrutura hierárquica de endereçamento, a alocação de endereços *multicast* pode ser complexa.

Os endereços *multicast* podem ser alocados estaticamente, usando um mecanismo chamado GLOP, ou dinamicamente, através da arquitetura MAAA (*Multicast Address Allocation Architecture*) [1].

III.3 Gerenciamento de Grupo

No IPv4, o protocolo IGMP (*Internet Group Management Protocol*) é usado pelas estações para informar ao primeiro roteador *multicast* os grupos dos quais ela participa. O IGMP é um protocolo ponto-a-ponto, da camada de rede e de escopo local. A construção da árvore de distribuição *multicast* se dá independentemente do IGMP. O protocolo possui ainda três versões.

III.3.1 O Protocolo IGMP, Versão 1

A versão 1 do IGMP foi especificada junto ao serviço IP Multicast [2]. No IGMPv1 só existem dois tipos de mensagens: relatório (*report*) e interrogação (*query*). Esta versão tornou-se obsoleta com a especificação da versão 2.

III.3.2 O Protocolo IGMP, Versão 2

A principal diferença da versão 2 [16] do IGMP para a versão 1 foi introdução da mensagem "deixar grupo" (*leave*), que acelera a reconfiguração da árvore de distribuição e evita o desperdício de recursos da rede. Na versão anterior, uma estação indicava a saída do grupo deixando de enviar mensagens "relatório". Com isso, a árvore de distribuição só era reconfigurada após o estouro do temporizador no roteador.

Na versão 2 existem dois subtipos de mensagem "interrogação": a "geral" (*general query*) e a de "grupo específico" (*group-specific query*). A primeira é usada pelos roteadores para descobrir quais grupos possuem membros em sua rede local. Já segunda é usada pelos roteadores para descobrir se um determinado grupo possui membros em sua rede local. O formato de todas as mensagens é o mesmo, Figura III.1.

0 1 2 3

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

tipo de mensagem	tempo máx. de resposta	soma de verificação
endereço do grupo		

Figura III.1: O formato das mensagens do protocolo IGMPv2.

O campo "endereço do grupo" da mensagem "interrogação geral", é preenchido com zeros. Já o mesmo campo das mensagens "interrogação de grupo específico" contém o endereço de um determinado grupo.

O funcionamento das duas versões do protocolo é similar e pode ser dividido em duas partes: o lado roteador e o lado estação. O roteador *multicast* local, ou de primeiro salto, envia periodicamente mensagens "interrogação", perguntando às estações conectadas à rede local sobre a existência de membros de algum grupo *multicast*. As mensagens "interrogação", são endereçadas ao grupo de "todas as estações" (*all-hosts group*), endereço `224.0.0.1`, com o campo TTL igual a 1 para que não sejam propagadas para fora da rede local.

Somente um roteador por rede local atua como "interrogador". Caso existam mais de um roteador *multicast* de primeiro salto por rede local, o "interrogador" será aquele que possui o menor endereço IP [17].

Cada estação que deseje escutar um grupo responde à interrogação com uma mensagem "relatório". É emitido um relatório para cada grupo no qual a estação está interessada. Vale ressaltar que um estação pode enviar um relatório informando o interesse em grupo *multicast*, a qualquer instante, sem que este tenha sido solicitado pelo roteador.

Para evitar a implosão do roteador *multicast* local, é usado um mecanismo de cancelamento de resposta. Um receptor não envia o seu relatório imediatamente após a recepção de uma interrogação, mas apenas após um intervalo de tempo aleatório, durante o qual ele escuta os relatórios emitidos por outras estações. Se os grupos nos quais o receptor tem interesse forem pedidos por outras estações, ele poderá suprimir o envio do seu relatório.

Quando uma estação deixa o grupo, ela envia uma mensagem "deixar grupo" ao grupo de "todos os roteadores", indicando explicitamente a sua desconexão. Ao receber uma mensagem "deixar grupo", o roteador *multicast* local envia uma mensagem de "interrogação de grupo específico". Caso o roteador não receba uma mensagem relatório para o determinado grupo em um certo tempo, ele assume que não há mais membros para aquele grupo, conectados à rede local.

III.3.3 O Protocolo IGMP, versão 3

A terceira versão do IGMP adiciona o suporte à "filtragem de fontes", isto é, as estações agora podem relatar o interesse na recepção de pacotes *multicast* enviados "apenas" por uma determinada fonte, ou enviados por todas as fontes "exceto" por uma determinada fonte. Esta informação pode ser usada pelos protocolos de roteamento *multicast* para evitar a distribuição de pacotes de fontes específicas em redes onde não existem receptores interessados [18].

Para que seja possível usar a filtragem de fontes, é necessário que a interface de serviço IP seja modificada. Um pedido IGMPv3, além do identificador do *socket*, da interface e do endereço do grupo *multicast*, deve conter o tipo de filtro, "INCLUDE" ou "EXCLUDE", e a lista de fontes onde deve ser aplicado o filtro. O IGMPv3 possui dois tipos de mensagens: interrogação e relatório. Para deixar um grupo, uma estação envia uma mensagem relatório bloqueando a recepção dos dados enviados para este. A função desta mensagem é a mesma da mensagem "deixar grupo" da versão 2 do IGMP. O IGMPv3 também suporta as mensagens das versões anteriores para manter a interoperabilidade.

O formato da mensagem interrogação é mostrado na Figura III.2. A principal diferença para as versões anteriores é que a mensagem pode conter uma lista de fontes de interesse.

O conteúdo deste campo só é diferente de zero quando o roteador de primeiro salto envia uma mensagem de interrogação específica para um determinado grupo e uma determinada fonte [18].

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

<table>
<tr><td colspan="3">tipo de mensagem</td><td>tempo máx. de resposta</td><td>soma de verificação</td></tr>
<tr><td colspan="5">endereço do grupo</td></tr>
<tr><td>reservado</td><td>S</td><td>robustez</td><td>intervalo de interrogação</td><td>número de fontes (n)</td></tr>
<tr><td colspan="5">endereço da fonte (1)
. . .
endereço da fonte (n)</td></tr>
</table>

Figura III.2: O formato da mensagem “interrogação” do protocolo IGMPv3.

A Figura III.3 mostra o formato da mensagem relatório. A principal diferença para as versões anteriores são os “registros de grupo”. Cada registro de grupo contém o tipo do registro, o número de fontes, o endereço do grupo *multicast* ao qual deve ser aplicada a filtragem de fontes e a lista de fontes. No fim das mensagens podem existir dados auxiliares, cujo comprimento é indicado pelo campo *comprimento dos dados auxiliares*. Todos os campos do registro de grupos são mostrados na Figura III.4

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

<table>
<tr><td>tipo de mensagem</td><td>reservado</td><td>soma de verificação</td></tr>
<tr><td colspan="2">reservado</td><td>número de registros de grupo (m)</td></tr>
<tr><td colspan="3">registro de grupo (1)
. . .
registro de grupo (m)</td></tr>
</table>

Figura III.3: O formato das mensagem “relatório” do protocolo IGMPv3.

Cada entrada da lista de fontes corresponde a um endereço IPv4 *unicast*. O campo “tipo de registro” indica o modo do filtro IGMPv3 que deve ser aplicado a todas as fontes listadas em um registro de grupo. O conteúdo do campo “tipo de registro” pode ser [18]:

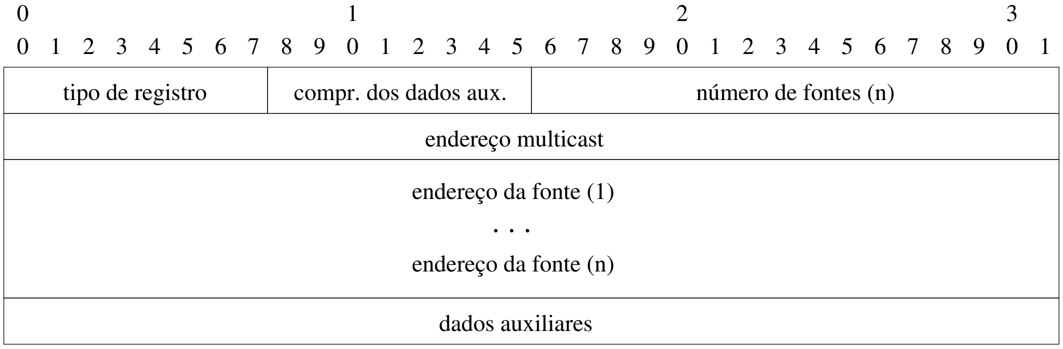


Figura III.4: O formato do campo "registro de grupo" da mensagem "relatório" do protocolo IGMPv3.

- MODE_IS_INCLUDE, indicando que os pacotes *multicast* endereçados ao grupo, especificado no registro de grupo, e provenientes das fontes listadas devem ser aceitos pela interface;
- MODE_IS_EXCLUDE, indicando que os pacotes *multicast* endereçados ao grupo, especificado no registro de grupo, e provenientes das fontes listadas devem ser rejeitados pela interface;
- CHANGE_TO_INCLUDE_MODE, indicando que, a partir de agora, os pacotes *multicast* endereçados ao grupo, especificado no registro de grupo, e provenientes das fontes listadas devem ser aceitos pela interface;
- CHANGE_TO_EXCLUDE_MODE, indicando que, a partir de agora, os pacotes *multicast* endereçados ao grupo, especificado no registro de grupo, e provenientes das fontes listadas devem ser rejeitados pela interface;
- ALLOW_NEW_SOURCES, indicando que as fontes listadas no registro de grupo devem ser adicionadas, no caso do modo do filtro ser INCLUDE, ou removidas, no caso do modo do filtro ser EXCLUDE, da lista de fontes já existente;
- BLOCK_OLD_SOURCES, indicando que as fontes listadas no registro de grupo devem ser adicionadas, no caso do modo do filtro ser EXCLUDE, ou removidas, no caso do modo do filtro ser INCLUDE, da lista de fontes já existente.

III.4 Roteamento

O IGMP é o protocolo responsável pelo diálogo entre as estações receptoras e o primeiro roteador *multicast*. Ele é um protocolo de escopo local. Para que o tráfego *multicast* seja encaminhado através de vários saltos, é necessária a criação de uma árvore de distribuição. A construção dessa árvore é a principal função de um protocolo de roteamento *multicast*.

Uma árvore de distribuição *multicast* pode ser construída de acordo com diferentes algoritmos. No entanto, os principais protocolos de roteamento usam dois tipos de árvores: árvores por fonte e árvores centradas, ou compartilhadas. Nas seções seguintes são descritos os algoritmos para a construção de árvores por fonte e árvores centradas.

III.4.1 Árvores de Distribuição por Fonte

Para cada fonte de dados *multicast* é necessária a criação e a manutenção de uma árvore por fonte. Dessa forma, a raiz da árvore é o próprio nó fonte [1].

Uma árvore por fonte pode ser construída através do algoritmo RPF (*Reverse Path Forwarding*). De forma simplificada, o algoritmo RPF funciona da seguinte maneira [17]:

1. ao receber um pacote *multicast* vindo da fonte S, o roteador R armazena a interface I, por onde recebeu o pacote;
2. se I é a interface que o roteador R usaria para enviar dados à fonte S, R deve retransmitir o pacote por todas as suas interfaces, exceto por I;
3. se o teste do item 2 for falso, o roteador R deve descartar o pacote.

O item 2 é chamado de teste RPF. Este teste indica qual a interface de um roteador será usada na árvore de distribuição de uma determinada fonte, ou seja a interface da saída para o caminho de menor custo. A Figura III.5 mostra um exemplo de árvore por fonte, construída com o algoritmo RPF. Cada enlace possui um custo associado. O nó A é a fonte e, inicialmente, envia pacotes *multicast* a todos os seus vizinhos, neste caso

B e D. Este nós recebem os pacotes *multicast* de A e aplicam o teste RPF. Como o resultado do teste é positivo, B e D retransmitem os pacotes vindos de A. Agora os pacotes *multicast* chegam aos nós C e E. Ambos aplicam o teste RPF. No caso de C, os pacotes chegam pela interface de saída para o caminho mais curto para A. Já em E, os pacotes *multicast* chegam por duas interfaces. A interface de saída para caminho de menor custo é selecionada.

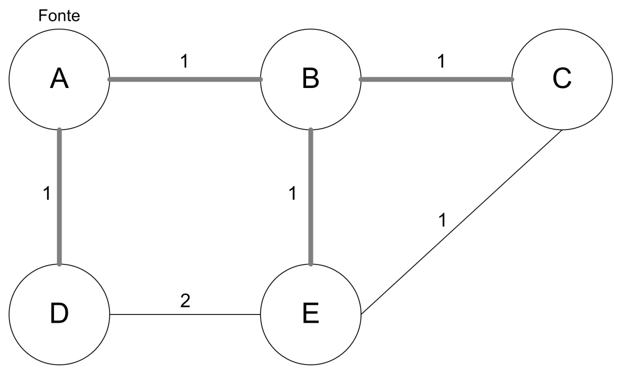


Figura III.5: Um exemplo de árvore RPF, onde o nó A é a fonte.

O algoritmo RPF, anteriormente descrito, pode ser melhorado. Para isto, basta que o roteador local R tenha conhecimento de que ele está no caminho mais curto entre um vizinho V e a fonte S, antes de retransmitir o pacote para o vizinho V. Se o resultado deste teste for verdadeiro, R encaminha o pacote para V. Porém, se o teste for falso, o pacote é descartado ainda no roteador local R. A informação necessária para a decisão de encaminhamento do pacote pode ser obtida facilmente, caso esteja sendo usado um protocolo de roteamento baseado no estado do enlace, uma vez que cada roteador mantém um mapa da topologia de todo o domínio de roteamento.

A principal vantagem do algoritmo RPF é garantir a distribuição eficiente dos dados, pois os pacotes *multicast* sempre seguem o caminho mais curto entre a fonte de dados e os receptores de um grupo. Isto é ideal para aplicações multimídias, que possuem restrições de tempo. Além disso, o algoritmo RPF é fácil de ser implementado, não requer que um roteador conheça toda a árvore de distribuição, nem requer um mecanismo especial para parar processo de envio de dados.

O problema do RPF é o envio de dados a todos os nós da rede, inclusive aqueles que não têm interesse neste tráfego. Para reduzir este problema, foram introduzidas mensagens de poda no algoritmo. A árvore construída por este algoritmo é chamada de TBT (*Truncated Broadcast Tree*).

No algoritmo RPF com poda, conhecido também como inundação-e-poda, o primeiro passo é a inundação da rede. Todos os roteadores recebem todos os pacotes *multicast*. Em seguida, um roteador "folha", que recebe um pacote *multicast* para o qual ele não possui um receptor interessado, envia uma mensagem de poda na direção contrária do fluxo. Dessa forma, os roteadores intermediários vão sendo informados da ausência de receptores interessados no grupo *multicast*, em partes da árvore, e deixam de enviar dados para estas regiões.

Apesar de evitar o tráfego em partes da árvore de distribuição, o algoritmo RPF com poda ainda necessita da inundação periódica de toda a rede, para que novos receptores possam se conectar à árvore. Além disso, todos os roteadores, mesmo aqueles que estão em áreas onde não existem receptores interessados, devem manter um estado para cada grupo *multicast*.

III.4.2 Árvores de Distribuição Centradas

Uma árvore de distribuição centrada é usada por várias fontes de dados. O nó raiz neste tipo de árvore é um nó arbitrário, idealmente localizado no "centro" da rede [1].

A árvore de distribuição é construída a partir do nó central. Os roteadores de primeiro salto enviam pedidos de conexão ao grupo para este nó central. Os roteadores atravessados pelos pedidos de conexão identificam a interface de entrada de cada pedido e armazenam essa informação para construir a árvore de distribuição.

A grande vantagem das árvores de distribuição centradas é o número reduzido de estados armazenados pelos roteadores. Apenas os roteadores que pertencem à árvore e a um dado grupo *multicast*, guardam estados para este grupo. Além disso, nenhum roteador mantém estado por fonte, o que aumenta a escalabilidade do algoritmo. Entretanto a escolha do nó central é o maior problema para as árvores de distribuição centradas. O algoritmo pode resultar na concentração do tráfego próximo ao nó central, uma vez que o tráfego de todas as fontes atravessa os mesmos enlaces ao se aproximar do centro [17].

III.4.3 Protocolos de Roteamento

A função de um protocolo de roteamento *multicast* é construir a árvore de distribuição de dados. Estes protocolos podem ser classificados em intra-domínio e inter-domínio. Os protocolos intra-domínio são responsáveis pelo roteamento dentro de um sistema autônomo. Já os protocolos inter-domínio são responsáveis pelo roteamento entre sistemas autônomos. Neste trabalho só serão abordados os protocolos de roteamento intra-domínio, visto que os testes realizados (ver Capítulo VI) envolvem apenas protocolos desta natureza.

O Protocolo DVMRP

O DVMRP (*Distance Vector Multicast Routing Protocol*) [17] foi o primeiro protocolo de roteamento *multicast* desenvolvido para a Internet. Ele usa vetores de distância, assim como o protocolo de roteamento *unicast* RIP (*Routing Information Protocol*), para construir rotas *unicast* para cada fonte de dados *multicast*. O DVMRP usa o algoritmo inundação-e-poda, descrito na Seção III.4.1 para construir árvores de distribuição TBT.

De acordo com o algoritmo do DVMRP, o primeiro datagrama IP para cada par (fonte,grupo) é enviado por toda a rede. Esse datagrama é limitado pelo campo TTL e, ao atingir roteadores folha que não possuem estações interessadas no grupo, dispara o envio de mensagens de poda na direção da fonte. As mensagens de poda constroem uma árvore particular com o caminho mais curto para cada fonte.

O uso das mensagens de poda indica a ausência de receptores interessados em um grupo *multicast* e não o interesse dos receptores. Portanto, é necessário inundar periodicamente toda a rede para informar a existência de novas fontes aos eventuais receptores em fontes podadas [17].

Para reduzir o tempo de reconfiguração da árvore de distribuição, o DVMRP possui um mecanismo para enxertar ramos podados. O mesmo roteador que enviou a mensagem de poda pode enviar mensagens de enxerto (*graft*) na direção da fonte, assim que descobrir membros para um determinado par (fonte,grupo) conectados a sua rede local.

O DVMRP, periodicamente, troca mensagens de atualização das tabelas de roteamento com roteadores *multicast* vizinhos. A tabela de roteamento não contém informações sobre os membros dos grupos *multicast*. Ela é composta pelo endereço da sub-rede fonte, a máscara para esta sub-rede, o roteador anterior na direção da sub-rede fonte e o campo TTL, usado como temporizador para a retirada de entradas da tabela. O campo "sub-rede fonte" contém o endereço de uma sub-rede que possua estações gerando tráfego *multicast*.

Além da tabela de roteamento, o DVMRP usa uma tabela de envio que contém os seguintes campos: sub-rede fonte, o grupo *multicast*, um campo "porta de entrada" para cada para (fonte,grupo) e um campo "portas de saída", através das quais datagramas destinados a um par (fonte,grupo) são enviados. No caso da tabela de envio, o campo "sub-rede fonte" corresponde à sub-rede que contém estações gerando dados *multicast* endereçados a grupos específicos. Uma sub-rede fonte pode conter fontes para vários grupos *multicast*.

A Figura III.6 mostra a construção da árvore de distribuição DVMRP para a fonte S, localizada na rede N_1. Os identificadores t_i em cada seta dessa figura ilustram a relação temporal entre os eventos. O roteador R_1 está diretamente conectado à rede N_1. Portanto, anuncia uma rota com métrica "1 salto" para seus vizinhos, R_2 e R_4. Tanto R_2 quanto R_4 receberam apenas um anúncio de rota para N_1 e usarão R_1 como próximo salto para N_1. Para comunicar ao roteador R_1 que R_2 e R_4 o usarão como "pai" na árvore *multicast*, o DVMRP usa um mecanismo de "*poison-reverse*" especial [1]. Os roteadores R_2 e R_4 enviam um vetor de distância com métrica igual a 33 (32 + métrica recebida) ao roteador R_1. O valor 32 corresponde ao infinito no DVMRP. Ao receber o vetor de

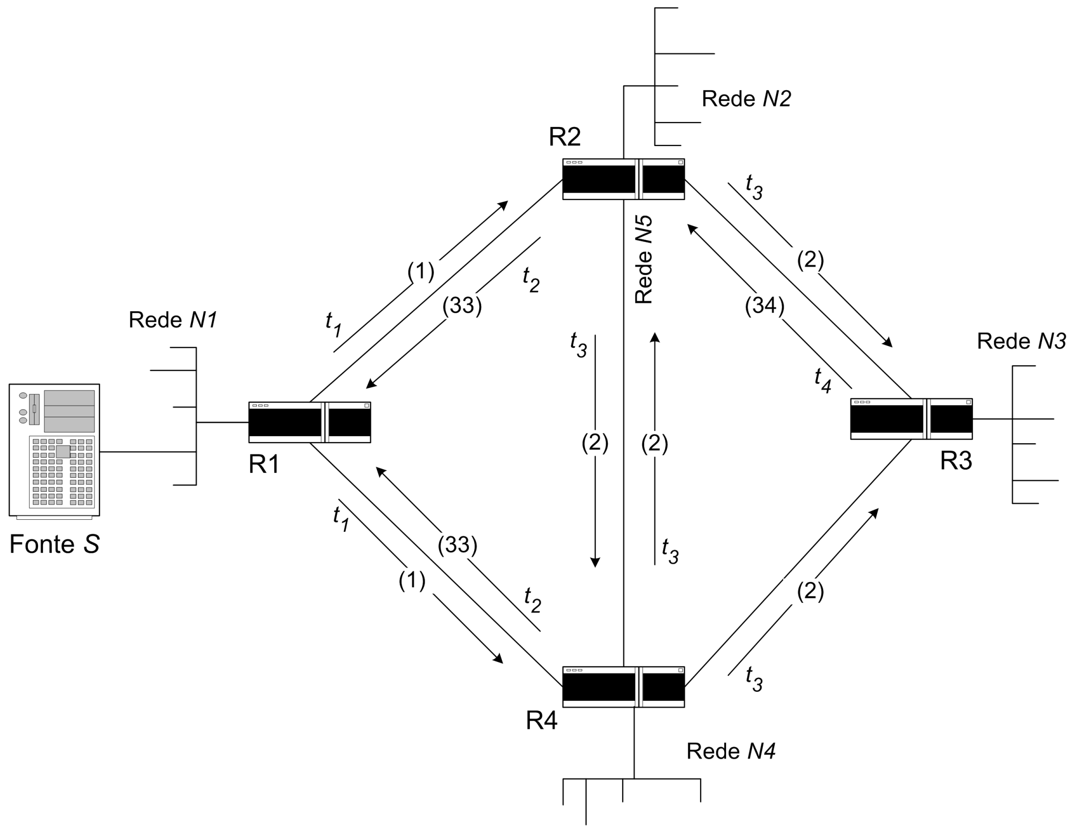


Figura III.6: Um exemplo de construção da árvore de distribuição do DVMRP.

distância com a métrica 33, R_1 marca, como "filhas", as interfaces que o conecta a R_2 e a R_4, construindo dois ramos da árvore. No passo seguinte, R_2 e R_4 enviam vetores de distância com métrica igual a 2 para alcançar N_1. Os vetores recebidos por R_2 e R_4 em suas interfaces conectadas à rede N_5 são ignorados, pois anunciam uma métrica para N_1 maior do que aqui eles possuem. O roteador R_3 recebe dois anúncios de rota para N_1, com a mesma métrica. Ele irá escolher a enviada pelo roteador de menor endereço IP, neste caso considera-se R_2. R_3 envia, então, a R_2 um vetor de distância com métrica 34, construindo mais um ramo da árvore TBT.

Supondo agora que existem receptores nas redes N_2 e N_3, interessados no conteúdo distribuído pela fonte S que envia dados para o grupo G. Inicialmente, a árvore TBT cobre todas as sub-redes da topologia, apresentada na Figura III.7. Logo, os dados enviados por S irão inundar toda a rede. Como na rede N_4 não existem receptores interessados no

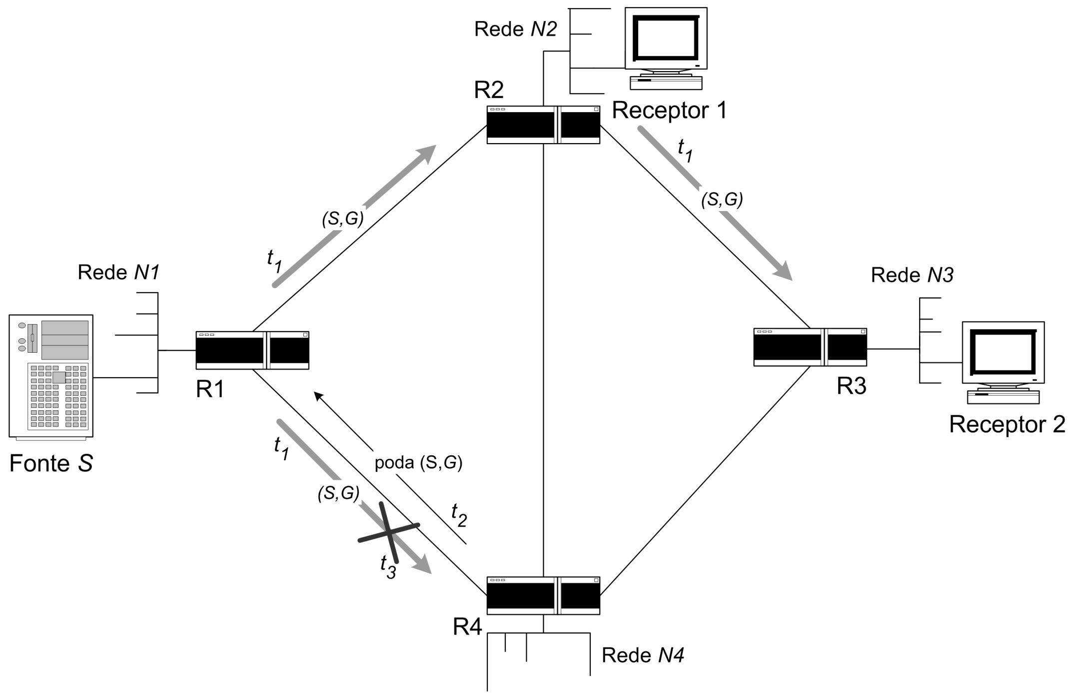


Figura III.7: Um exemplo do funcionamento do DVMRP.

grupo G, R_4 enviará uma mensagem de poda ao roteador R_1. Isto fará com que o tráfego cesse no enlace entre os roteadores R_1 e R_4.

O Protocolo PIM

O PIM (*Protocol Independent Multicast*) é o protocolo de roteamento *multicast* mais popular da Internet. Ele é opera em dois modos: o denso (PIM-DM) e o esparso (PIM-SM). Os dois modos de operação estão associados à distribuição dos receptores na rede. Um grupo é considerado denso quando a probabilidade de uma determinada área ter pelo menos um membro do grupo é alta. Do contrário, se a probabilidade for pequena, o grupo é considerado esparso.

O PIM-DM (PIM - *Dense Mode*) [19] foi desenvolvido para ser usado em redes onde os membros de um grupo *multicast* estão densamente distribuídos. Seu funcionamento é similar ao do DVMRP. O PIM-DM também usa a inundação periódica da rede com mensagens de poda. Porém, ao contrário do DVMRP que implementa o próprio mecanismo de roteamento *unicast*, o PIM-DM utiliza as informações do protocolo *unicast* em uso na rede para construir a sua árvore de distribuição.

O PIM-SM (PIM - *Sparse Mode*) [20] foi projetado para o roteamento em grande escala, em que a distribuição dos receptores é esparsa. O PIM-SM constrói árvores centradas e unidirecionais a partir de mensagens de inscrição (*join*). Os dados são encapsulados em *unicast* e enviados pela fonte a um nó de "ponto de encontro" (RP - *Rendez-vous Point*). Em seguida, o RP distribui os dados na árvore *multicast*. Para fontes com alta taxa de transmissão, é possível mudar para uma árvore de distribuição por fonte. Isto torna menos crítica a localização do nó central (RP).

O Protocolo PIM-DM

Tanto o PIM, modo denso, e o DVMRP usam o mesmo algoritmo para construir suas respectivas árvores de distribuição, porém, o PIM-DM é um protocolo mais simples do que o DVMRP. O PIM-DM se baseia na existência de um protocolo de roteamento *unicast* para se adaptar a mudanças de topologia na rede, mas independe dos mecanismos do protocolo específico. Já o DVMRP implementa um protocolo de roteamento *unicast*, com troca de mensagens para construção e atualização das tabelas de rotas.

Além disso, O PIM-DM não calcula um conjunto de interfaces "filhas", como faz o DVMRP, para construir sua árvore de distribuição. Ele simplesmente envia o tráfego *multicast* em todas as interfaces no sentido descendente do fluxo, até que mensagens de poda explícita sejam recebidas.

Se um roteador PIM, modo denso, recebe um pacote *multicast* de um determinado par (fonte,grupo), ele verifica a sua tabela de roteamento *unicast* e também se a interface de chegada do pacote é uma das interfaces usadas para transmitir pacotes *unicast* para aquela fonte. Se o resultado do teste for negativo, o pacote *multicast* é descartado e uma mensagem de poda para o par (fonte,grupo) é enviada pela interface de chegada do pacote. Porém, se o resultado do teste for positivo, o roteador encaminha uma cópia do pacote *multicast* em todas as suas interfaces, nas quais ainda não tenha recebido uma mensagem de poda para o par (fonte,grupo). Caso todas as interfaces tenham sido “podadas”, o roteador descarta o pacote *multicast* e envia uma mensagem de poda (fonte,grupo) pela interface de chegada do pacote.

A Figura III.8 mostra o funcionamento do processo de inundação e poda do protocolo PIM-DM. Existe uma fonte de tráfego *multicast* S, localizada na rede N_1, que inunda toda a rede. Existem também três receptores membros do grupo G, para os quais o tráfego de S é encaminhado: um localizado na rede N_2, outro na rede N_3 e o último da rede N_4. Cada roteador aplica o teste, anteriormente descrito, e em caso positivo reenvia o pacote *multicast* em todas as suas interfaces, onde existem vizinhos PIM-DM. Desta forma, todos os enlaces da rede são usados, o que não ocorre no DVMRP. Além disso, alguns pacotes podem ser transportados em um enlace nos dois sentidos, como ocorre, neste exemplo, entre os roteadores R_2 e R_4. A duplicação de pacotes é o preço pago pelo PIM-DM para eliminar a dependência do protocolo de roteamento *unicast*.

O roteador R_3 possui dois caminhos para alcançar a fonte S. Neste caso, R_2 envia a R_3 e a R_4 mensagens de acordo (*assert*). Os roteadores R_3 e R_4 também enviam mensagens de acordo para R_2 em resposta a mensagem recebida. Uma mensagem de acordo contém a métrica que os roteadores, que a emitiram, possuem para chegar até a fonte. Um roteador que recebe uma mensagem de acordo compara a métrica contida nesta mensagem com a sua métrica atingir a fonte. Se a sua métrica for maior, o roteador deixa de enviar pacotes para o par (S,G) na interface na qual ele recebeu o “acordo”. Se os roteadores possuem a mesma métrica, o roteador de menor endereço IP continua a enviar os pacotes *multicast*. Supondo que as métricas de R_2 e R_4 sejam iguais, mas que R_3 possua um endereço IP maior, o roteador R_3 enviará uma mensagem de poda na interface que o conecta a R_2.

Os roteadores R_2 e R_4 também enviam podas nas interfaces que não foram indicadas pelo teste RPF, já que receberam tráfego vindo de S por mais de uma interface durante a inundação inicial.

A inundação do PIM-DM gera uma árvore de distribuição de menor caminho entre os receptores e a fonte. Além disso, cada roteador, mesmo que não esteja recebendo tráfego

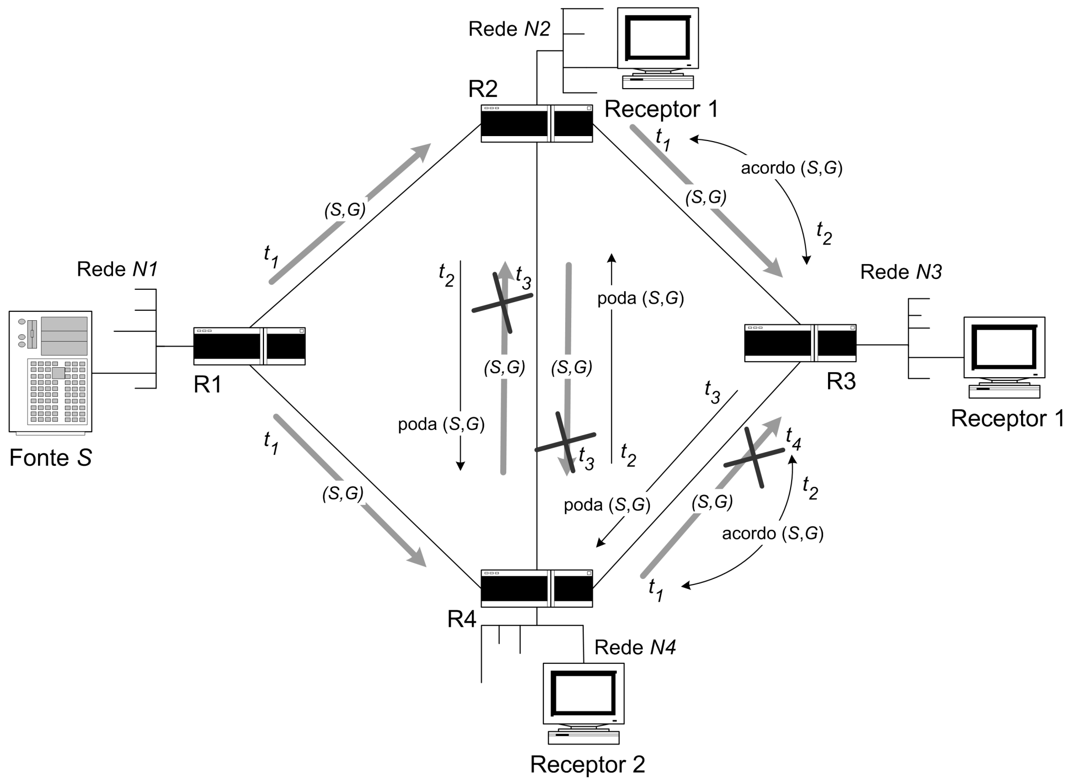


Figura III.8: Um exemplo do funcionamento do PIM-DM.

vindo de S, mantém um estado para o par (S,G). Este estado continuará armazenado enquanto S está transmitindo. As podas também são estados voláteis, que expiram de acordo com o temporizador configurado no roteador. Portanto, a rede é inundada a cada vez que este temporizador expira. Por isso, o PIM-DM só deve ser usado em redes com receptores densamente distribuídos.

O Protocolo PIM-SM

O protocolo PIM-SM foi projetado para limitar o tráfego *multicast* de forma que somente os roteadores interessados em receber tráfego de um determinado grupo o enxerguem. Isto é ideal para cenários em que os receptores se concentram em partes da rede e para cenários em que o número de receptores é pequeno.

No PIM, modo esparso, os roteadores com membros do grupo diretamente conectados, se conectam à árvore de distribuição através do envio de mensagens explícitas de inscrição (*join*). Roteadores que fazem parte do caminho entre um membro, conectado a outro roteador, e a fonte também enviam mensagens de inscrição para se conectarem a árvore.

O uso de mensagens de inscrição explícitas é a principal diferença do protocolo PIM de modo esparso para o de modo denso. Enquanto no PIM-DM, o tráfego é enviado até que mensagens explícitas de poda sejam recebidas, no PIM-SM o tráfego é bloqueado até que seja explicitamente solicitado.

Outra diferença entre os dois modos de operação do PIM é o tipo de árvore construída. No PIM-SM, a árvore de distribuição é centrada, ou compartilhada. Os nós centrais da rede são usados como raiz da árvore e são chamados de pontos de encontro (RP – *Rendez-vous Point*), por serem o local onde os receptores "encontram" as fontes. Para cada grupo são associados um, ou mais, pontos de encontro, através de um mecanismo de eleição [20]. Entretanto, para cada grupo só existe um RP ativo.

Quando um receptor deseja se tornar membro de um grupo, ele deve informar o seu interesse ao seu roteador local. O responsável por esta sinalização é o protocolo IGMP. Para se conectar à árvore de distribuição, o roteador local envia uma mensagem de inscrição ao RP do grupo. O roteador local, neste caso, é chamado de designado (DR – *Designated Router*). Sendo assim, a árvore de distribuição é construída, de forma explícita, através do envio de inscrições ao RP. Todas as fontes também devem se registrar no RP. Cada fonte envia apenas uma cópia de cada pacote *multicast* para os receptores conectados ao RP.

Cada roteador, mesmo que não faça parte de um grupo, deve manter a lista dos RPs, contendo o mapeamento entre os grupos e os RPs associados. Dessa forma, no PIM-SM, os roteadores devem manter algum estado mesmo antes da chegada dos pacotes *multicast*. Já no PIM-DM, os roteadores não guardam nenhum estado até a chegada de pacotes

As árvores compartilhadas podem apresentar atraso maior do que as árvores por fonte, uma vez que o tráfego tende a se concentrar nos enlaces próximos ao RP. Para minimizar os problemas causados por uma má localização do RP, o PIM-SM permite a troca da árvore centrada por uma árvore por fonte. Esta troca pode ser configurada de acordo com a taxa de envio de dados de uma determinada fonte S. A partir de um valor limite da taxa de dados recebidos da fonte S, os roteadores designados, no qual estão conectados os receptores, podem decidir pela troca do tipo da árvore. Uma vez decidida a troca, os roteadores designados emitem mensagens de inscrição para o par (fonte,grupo) na direção de S. As mensagens de inscrição são encaminhadas, nó a nó, até chegarem ao roteador de primeiro salto, no qual está conectada a fonte S. É importante ressaltar que os roteadores designados não param de enviar mensagens de inscrição ao RP, pois isso faria com que os receptores deixassem de receber tráfego de outras fontes ativas. Entretanto, para evitar que o tráfego gerado por S seja recebido pelo receptor através das duas árvores, o roteador envia ao RP uma mensagem de poda especial (RP–*bit–prune*) para o par (fonte,grupo).

A Figura III.9 ilustra o funcionamento do PIM-SM, usando uma árvore de distribuição compartilhada. A fonte S está localizada na rede N_1 e envia dados para o grupo G. Os receptores que fazem parte de G estão localizados nas redes N_3 e N_4. Considerando que o RP associado ao grupo G é o roteador R_2, os roteadores designados R_3 e R_4 enviam inscrições na direção de R_2, construindo assim a árvore de distribuição. Por sua vez, o roteador R_1 envia uma mensagem de registro ao RP, relativa à fonte S que está ativa. Na realidade, R_1 encapsula os dados enviados por S dentro das mensagens de registro, que enviadas são enviadas ao RP em *unicast*. Quando o RP recebe as mensagens de registro, ele desencapsula os dados *multicast* e os envia sobre a árvore compartilhada [1].

Supondo que a taxa de dados enviados por S ultrapasse o valor limite configurado em R_3 e R_4, os dois roteadores decidirão pela troca da árvore de distribuição. Esta situação é descrita na Figura III.10. Os roteadores designados R_3 e R_4 enviam mensagens de inscrição par o par (S,G), à fonte S. A cada salto atravessado pelas mensagens de inscrição, é criado um estado para o par (S,G). R_3 e R_4 também enviam mensagens de poda especial ao RP, no caso R_2. Dessa forma, a fonte S passa a utilizar uma árvore por fonte de caminho mais curto para enviar os pacotes *multicast*.

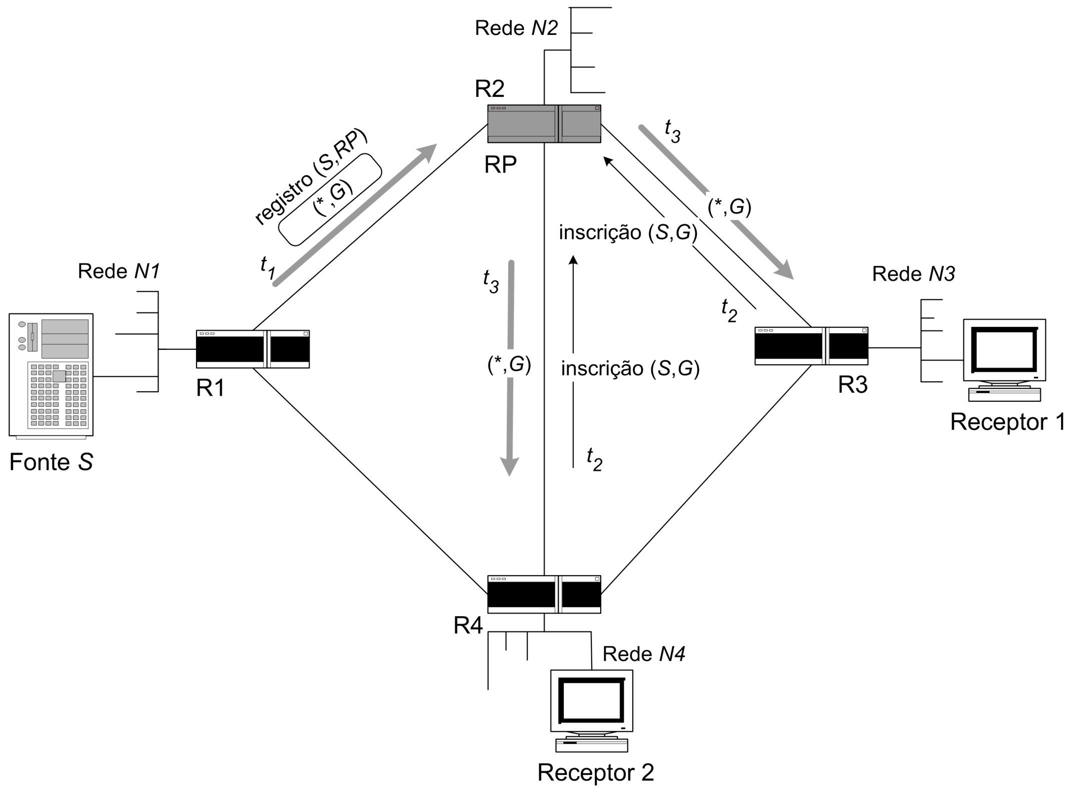


Figura III.9: Um exemplo do funcionamento do PIM-SM, com árvore compartilhada.

III.5 IP Multicast na Internet

O serviço IP Multicast começou a ser implantado experimentalmente na Internet com o desenvolvimento do MBone (*Multicast Backbone*) [5], em 1992.

O MBone é uma rede virtual construída sobre a topologia da Internet para a realização de experimentos envolvendo aplicações *multicast*. Como somente alguns roteadores da Internet implementam o serviço *multicast*, o MBone é constituído por "ilhas *multicast*" interconectadas por túneis. Uma ilha *multicast* é uma área com suporte ao serviço IP Multicast. O uso dos túneis possibilitou a rápida implementação de uma infra-estrutura *multicast* experimental. Todo o tráfego *multicast* é encapsulado em pacotes *unicast*, usando encapsulamento IP dentro de IP. Dessa forma, os pacotes *multicast* são encaminhados por roteadores *unicast* até atingirem a ilha *multicast* de destino. Inicialmente, o DVMRP era o protocolo de roteamento usado no interior de cada ilha *multicast*. Hoje em dia, o

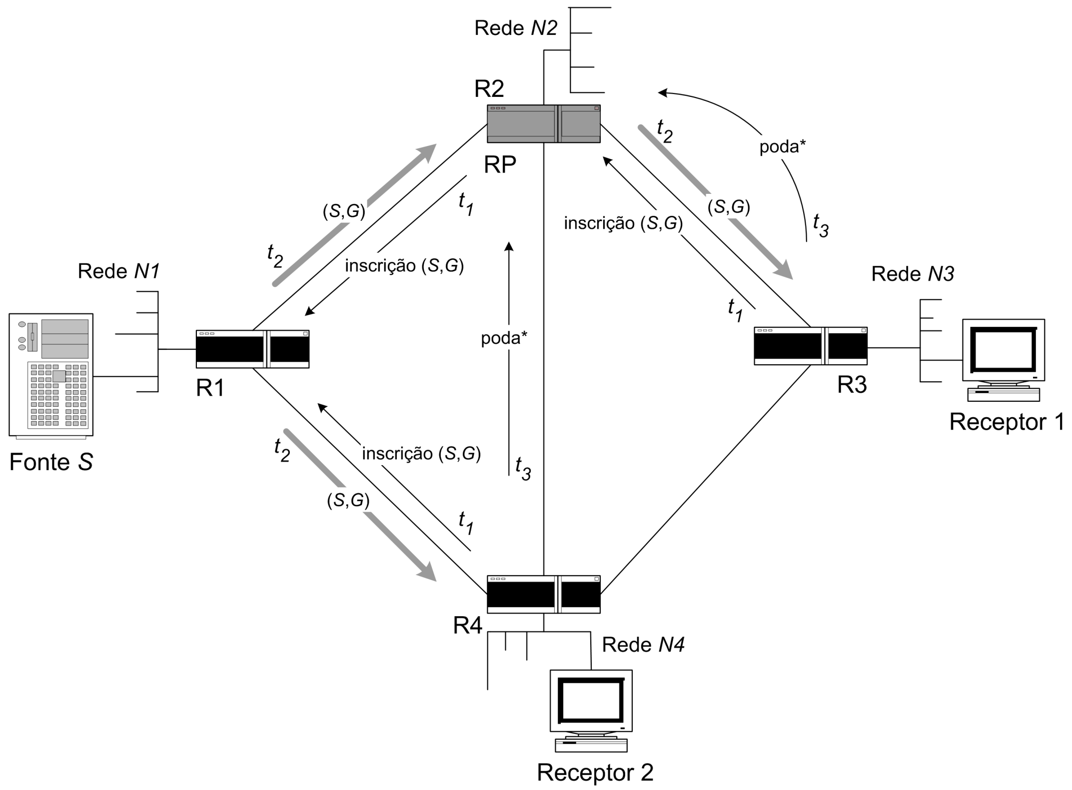


Figura III.10: Um exemplo do funcionamento do PIM-SM com árvore por fonte.

MBone continua operacional, mas os roteadores da Internet começam a oferecer suporte nativo ao serviço *multicast*.

A arquitetura atual do serviço IP Multicast é baseada no conjunto de protocolos IGMP/PIM/MBGP/MSDP, como mostra a Figura III.11. O IGMP é o protocolo responsável pelo diálogo dos receptores com os roteadores de último salto. O PIM é um protocolo de roteamento bastante flexível, que pode ser usado tanto em redes pequenas com muitos receptores, quanto em grandes redes com receptores distribuídos esparsamente distribuídos. Os protocolos MBGP (*Multiprotocol Extensions for BGP-4*) [17] e MSDP (*Multicast Source Discovery Protocol*) [17] representam uma solução temporária para o roteamento inter-domínio. Cada domínio rodando o protocolo PIM-SM possui um conjunto próprio de RPs para todos os seus grupos. Os RPs de um domínio estão conectados. O MSDP foi desenvolvido para conectar RPs em diferentes domínios, o que possibilita a distribuição do mapeamento de cada RP entre diferentes domínios. O MBGP permite que o protocolo de

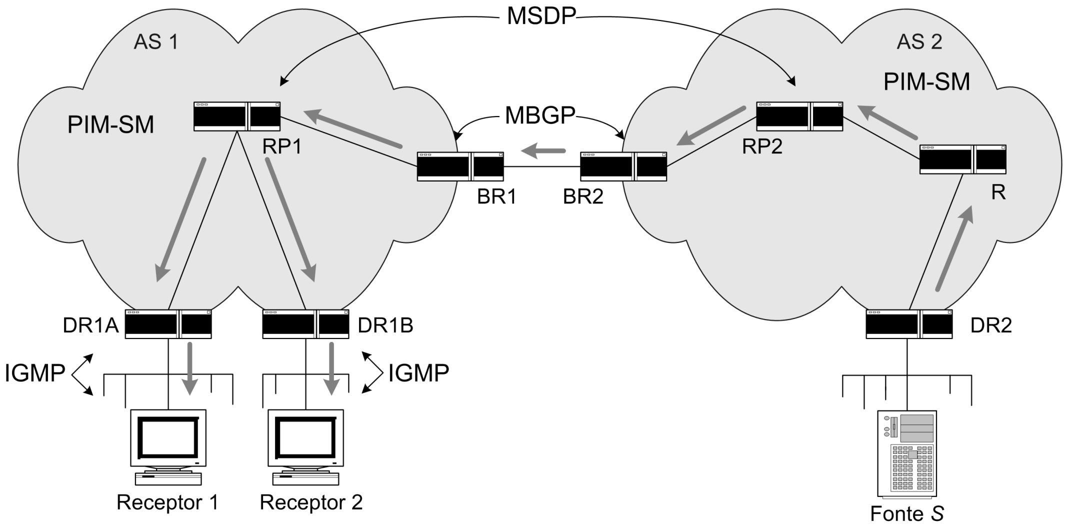


Figura III.11: A atual arquitetura do serviço IP Multicast na Internet.

roteamento *unicast* inter-domínio mais usado na Internet, o BGP (*Border Gateway Protocol*) [17], mantenha múltiplas tabelas de roteamento para diferentes protocolos. Dessa forma, os roteadores podem tratar diferenciadamente os pacotes *multicast* e os pacotes *unicast*.

Em um futuro próximo, o protocolo BGMP (*Border Gateway Multicast Protocol*), a versão *multicast* do BGP, irá substituir o conjunto MSDP/MBGP [17].

III.6 Problemas do Serviço IP Multicast

Existem diversos problemas associados à arquitetura e ao modelo de serviço do IP Multicast, que dificultam e retardam a sua implantação com um serviço nativo na Internet.

A alocação de endereços é um dos problemas mais graves do IP Multicast. Para que um grupo *multicast* seja criado, basta que uma fonte comece a enviar dados para um endereço IP da faixa `224/4`. Não há qualquer controle ou hierarquia sobre a alocação de endereços, uma vez que os endereços não foram delegados por instituições regulamentadoras, como o IANA. Com isso, existe a possibilidade da colisão de endereços e, conseqüentemente, da interferência entre aplicações. O endereçamento GLOP e a arquitetura MAAA tentam solucionar este problema, mas são complexos e, em muitos casos, comprometem a escalabilidade do serviço.

Outro problema é a falta de controle de acesso aos grupos *multicast*. É permitido que uma estação que não pertença ao grupo envie dados para este. Além disso, um receptor não pode especificar de quais fontes ele deseja receber tráfego, quando esta se inscreve em um dado grupo. Também não há como garantir que uma fonte não usará o mesmo endereço *multicast*, já alocado por outra fonte para transmitir seus dados.

No IP Multicast, a identidade e o número dos participantes de um grupo não são conhecidos, o que dificulta a adoção de mecanismos de segurança e tarifação de serviços. Atualmente, os grandes provedores de serviço da Internet firmam seus contratos com base na taxa de envio de dados do cliente. Por contrato, para uma taxa de envio do cliente igual a R, os provedores de serviço devem garantir uma taxa de entrega de $R \times n$, com $0 < n \leq 1$. O serviço IP Multicast viola este tipo de acordo, uma vez que a taxa de entrega de dados pode ser maior do que a taxa de envio do cliente, quando um grupo possui mais de um receptor.

A maior parte destes problemas é solucionada pelo serviço de difusão com fonte específica, o SSM (*Source-Specific Multicast*).

Capítulo IV

O Serviço SSM

O serviço IP Multicast não possui nenhum mecanismo de controle de acesso de estações ao grupo e requer um sistema de alocação global de endereços para evitar a colisão entre aplicações. Estas duas características dificultam a implantação do serviço em larga escala na Internet.

Uma alternativa ao IP Multicast é o serviço de difusão com fonte específica, ou serviço SSM (*Source-Specific Multicast*) [3]. Este novo serviço é uma simplificação do IP Multicast. O SSM reduz a conversação *multicast* para um-para-muitos já que usa o conceito de *canal multicast*. Um canal possui apenas uma fonte e zero ou mais assinantes e é identificado por uma par de endereços (fonte,destino).

Neste capítulo será apresentado o modelo de serviço SSM. Também será descrito o funcionamento do protocolo de roteamento usado pelo SSM, assim como os requisitos para a implantação deste serviço na Internet.

IV.1 O Modelo de Serviço SSM

O serviço SSM é uma simplificação do modelo IP Multicast. No SSM, a conversação *multicast* é reduzida para um-para-muitos. Esta redução é válida, pois a maioria das apli-

Tabela IV.1: Diferenças entre os modelos de serviço IP Multicast e SSM.

	IP Multicast	SSM
Abstração	grupo	canal
Identificador	endereço G	par de endereços (S,G)
Operação dos Receptores	"inscrever" e "deixar"	"assinar" e "cancelar"

cações de grupo conhecidas atualmente possui uma só fonte. Além disso, nas aplicações conhecidas com múltiplas fontes, o número de fontes é reduzido, o que permite o uso de múltiplos canais de fonte específica. Um exemplo de aplicação de fonte única é a distribuição de TV na Internet. Já a videoconferência é exemplo de aplicação com múltiplas fontes.

Para reduzir a conversação *multicast* para um-para-muitos, o SSM utiliza o conceito de *canal* proposto pelo protocolo EXPRESS (*EXPlicitly REquested Single-Source multicast*) [4]. Um *canal* é um serviço de entrega de datagramas identificado por um par de endereços (S,G), onde S é o endereço da fonte e G o endereço de destino do canal. Somente a fonte S pode enviar dados para o canal (S,G).

Um assinante solicita a recepção dos dados enviados a um canal, especificando explicitamente os endereços S e G, respectivamente, da fonte e do destino. Já a fonte envia os dados para o canal, especificando apenas o endereço G de destino do canal. A camada de rede garante que todos os datagramas enviados pela a fonte S para o grupo G são entregues a todos os assinantes do canal (S,G), com a mesma confiabilidade de um datagrama *unicast* enviado por S [4].

No IP Multicast, um receptor pode se "inscrever" e "deixar" um grupo. Já no SSM um receptor pode "assinar" e "cancelar a assinatura" de um canal. A Tabela IV.1 resume as diferenças entre os serviços IP Multicast e SSM.

Para assinar um canal, o receptor tem que usar um protocolo de gerenciamento de grupo, da mesma forma como faz para se inscrever em um grupo *multicast*. Porém, o protocolo de gerenciamento deve ser capaz de solicitar o recebimento dos dados enviados por apenas uma única fonte a um endereço de destino. Só a versão 3 do protocolo IGMP implementa um mecanismo de filtragem de fontes, descrito na Seção III.3.3.

Atualmente, o serviço SSM é implementado pelo IGMPv3 e pelo protocolo de roteamento PIM-SSM. O PIM-SSM é uma versão simplificada do protocolo PIM-SM, com novas regras de tratamento de mensagens e com a capacidade de interpretar corretamente as mensagens IGMPv3 enviadas pelas estações. Seu funcionamento é descrito na Seção IV.3.

Por ser mais simples e eliminar algumas das limitações da arquitetura de distribuição *multicast* PIM-SM/MBGP/MSDP, o modelo SSM foi bem aceito pela comunidade científica. O mesmo ocorreu com setores da indústria, já que o SSM tem o potencial de acelerar a implantação do serviço *multicast* na Internet. Desta forma, a padronização do serviço SSM no IETF está sendo concluída em um prazo mais curto que a média dos protocolos da Internet.

IV.2 Endereçamento

Um canal usado no serviço SSM é identificado por um par de endereços (fonte,destino). A fonte é identificada pelo endereço IP da única estação que envia dados ao canal. Já o endereço de destino do canal é alocado dinamicamente pela fonte e pertence à faixa `232/8`.

Essa faixa de endereços, de `232.0.0.0` a `232.255.255.255`, foi reservada pelo IANA, para o uso exclusivo do serviço SSM. Dessa forma é possível a coexistência dos serviços SSM e IP Multicast, pois, de acordo com faixa do endereço de grupo, o pacote *multicast* poderá ser tratado de forma diferenciada pelos roteadores. O IANA também definiu que o endereço `232.0.0.0` não pode ser atribuído a nenhum endereço de destino do canal.

No serviço SSM, não é necessário um mecanismo global de alocação de endereços, uma vez que um canal é identificado por um par (fonte,destino), em que o endereço IP da fonte é único em toda a rede. Cada fonte é responsável, localmente, por evitar a colisão de endereços para os vários canais criados por ela. Uma fonte pode ter até 2^{24} canais associados.

É importante ressaltar que qualquer estação pode enviar dados para um endereço de destino da faixa reservada, porém somente os assinantes do respectivo canal (fonte,destino) irão receber o tráfego gerado pela fonte.

IV.3 O Protocolo PIM-SSM

O PIM-SSM (*Protocol Independent Multicast – Source-Specific Multicast*) [21, 22] é uma versão simplificada do protocolo PIM-SM. O PIM-SSM constrói apenas árvores de distribuição por fonte. Para cada canal *multicast*, o protocolo considera que existe apenas uma fonte associada. Como o PIM-SSM não utiliza árvores compartilhadas, não é necessária a configuração de roteadores de ponto de encontro (RPs) para gerenciar a árvore nem de mecanismos de mapeamento de endereços *multicast* para RPs.

A coexistência dos serviços SSM e IP Multicast é possível. Para isso, basta que os roteadores tratem diferenciadamente cada grupo, de acordo com o seu endereço *multicast*. Se o endereço do grupo pertence à faixa `232/8`, reservada para o SSM, os pacotes são roteados pelo PIM-SSM. Se o endereço do grupo pertence ao restante da faixa `224/4`, os pacotes são tratados pelo protocolo PIM-SM.

O funcionamento do PIM-SSM é similar ao do protocolo PIM-SM, quando este usa árvores por fontes. O protocolo IGMPv3 é o responsável por sinalizar o interesse de um receptor em um par (fonte,destino) ao roteador local. Se o endereço do destino pertence a faixa exclusiva do serviço SSM, o roteador local, também chamado de roteador designado (DR), envia uma mensagem de assinatura na direção da fonte para se conectar à árvore de distribuição, correspondente ao canal (fonte,destino). Porém, se o endereço estiver fora da faixa SSM, o roteador designado enviará a mensagem de inscrição ao grupo em direção ao roteador RP e se conectará a árvore compartilhada.

A Figura IV.1 ilustra o funcionamento do PIM-SSM. A fonte S está localizada na rede N_1 e envia dados para o destino G. Os assinantes do canal (S,G) nas redes N_3 e N_4. Os roteadores designados R_3 e R_4 enviam, então, mensagens de assinatura para o canal (S,G), diretamente à fonte S. A cada salto atravessado pelas mensagens de assinatura, é criado um estado para o par (S,G). Dessa forma, é construída uma árvore de distribuição por fonte de caminho mais curto entre a fonte e os assinantes.

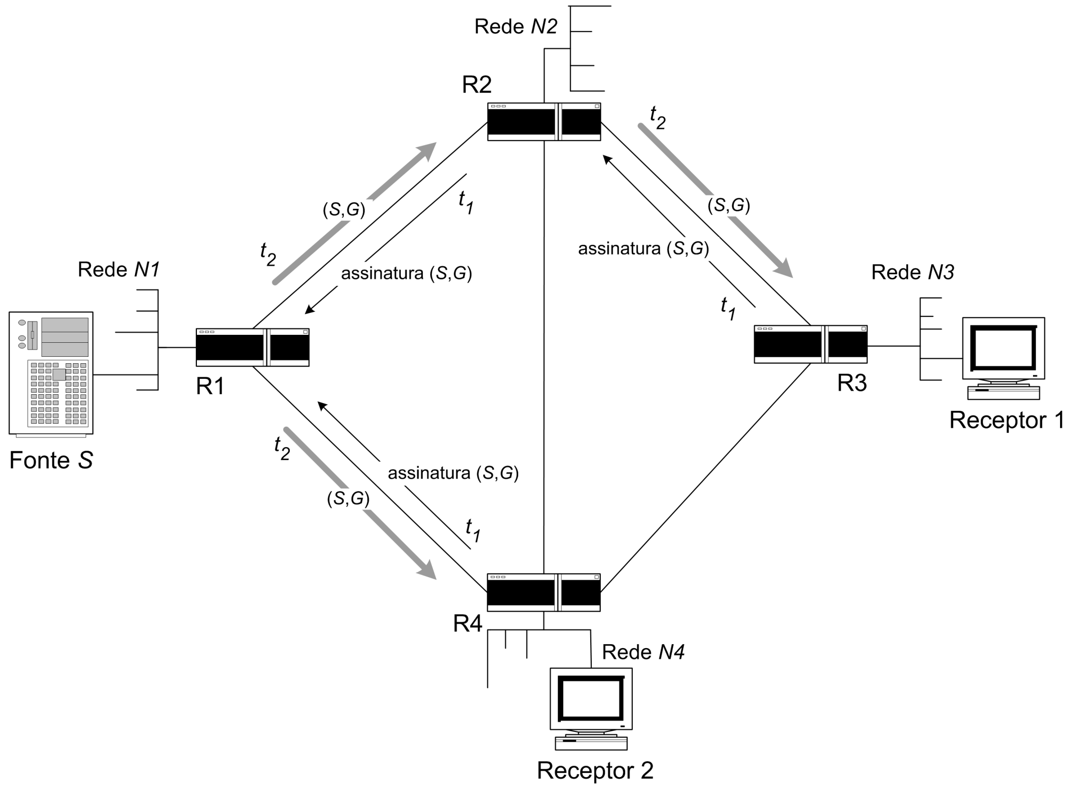


Figura IV.1: Um exemplo do funcionamento do PIM-SSM.

Para que os serviços IP Multicast e SSM possam coexistir, é necessário que protocolo PIM-SM seja modificado. O roteador RP não pode aceitar inscrições na faixa exclusiva SSM e nem deve se candidatar no processo de eleição de "ponto de encontro" para endereços nessa faixa. Os roteadores designados também não devem enviar mensagens de inscrição ao RP, caso o endereço do grupo pertença à faixa de endereços SSM.

O uso de mecanismos de descoberta de fontes no inter-domínio, como o protocolo MSDP, não são necessários, uma vez que o PIM-SSM não usa roteadores RP. O protocolo

PIM-SSM supõe que os receptores potenciais conhecem o endereço da fonte, sendo assim capazes de gerar pedidos IGMPv3 contendo este endereço. O mecanismo de descoberta de fontes utilizado não está definido pelo PIM-SSM. A descoberta de fontes pode ser realizada através de serviços como o correio eletrônico, publicação em sítios da Internet, ou pelo protocolo de anúncio de sessão, SAP (*Session Announcement Protocol*) [23].

IV.4 Benefícios do Serviço SSM

O modelo de serviço SSM soluciona a maior parte dos problemas apresentados pelo IP Multicast.

O serviço de difusão com fonte específica elimina o possibilidade de interferência entre aplicações, devido à colisões de endereços. O uso simultâneo de um endereço de destino SSM por duas, ou mais, fontes é possível, uma vez que cada canal é identificado por um par de endereços (fonte,destino).

Um mecanismo de controle de acesso também é oferecido pelo SSM. Quando um receptor assina um canal (S,G), ele recebe apenas os dados enviados pela fonte S. Da mesma maneira, quando uma fonte escolhe um canal para transmitir, existe a garantia de que nenhuma outra fonte usará aquele canal. Isto dificulta a ação maliciosa de estações, que visam prejudicar a transmissão em um canal *multicast*.

Mecanismos de autenticação e de contagem do número de participantes também podem ser implementados, pois os pedidos de assinatura de um canal são encaminhados diretamente à fonte pelos roteadores.

Como o SSM requer apenas árvores de distribuição por fonte, não existe a necessidade da configuração de roteadores de ponto de encontro (RP) e nem do uso do protocolo MSDP para a descoberta de fontes no inter-domínio. O protocolo de roteamento inter-domínio MBGP não precisa ser modificado para o serviço SSM [21]. Isto simplifica a arquitetura do serviço.

IV.5 Distribuição de TV sobre o Serviço SSM

O serviço SSM foi desenvolvido para a difusão de aplicações de grupo com uma fonte única e múltiplos receptores. Um dos exemplos mais difundidos deste tipo de aplicação é a distribuição de televisão.

O uso do protocolo PIM-SSM para a aplicação de distribuição de TV é vantajoso se comparado com o protocolo PIM-SM.

A grande vantagem do PIM-SSM para a TV na Internet é a construção de árvores de distribuição por fonte de menor caminho entre a fonte e os receptores. Este tipo de árvore proporciona menores atrasos entre a fonte e cada um dos assinantes do canal *multicast*, quando comparada a árvore compartilhada construída pelo PIM-SM. Isto é essencial para a transmissão de dados em tempo real.

É importante ressaltar que o PIM-SM também pode usar árvores por fonte para fontes com altas taxas de envio de dados. Entretanto, a conexão a um canal *multicast* não pode ser feita sem a passagem obrigatória por uma árvore compartilhada. Isto proporciona uma maior latência de conexão ao canal, o que significa que o início da recepção de dados *multicast* será mais lenta no PIM-SM que no PIM-SSM.

No PIM-SSM, o endereço da fonte deve ser conhecido para que a aplicação se inscreva no canal *multicast*. Para a distribuição de TV isto não é problema, uma vez que o endereço da fonte deve ser bem conhecido. Um mecanismo de divulgação do endereço da fonte e do endereço de destino que identificam um canal *multicast* deverá ser o responsável pelo controle de acesso ao canal. O identificador do canal só seria informado aos seus assinantes.

O serviço SSM também proporciona uma maior proteção para aplicações de fonte única, como a TV na Internet. De acordo com o modelo de serviço SSM, apenas a fonte S pode enviar dados ao endereço de destino G do canal (S,G). No modelo de serviço IP Multicast, um grupo *multicast* G é definido como uma conversação muitos-para-muitos, em qualquer estação pode enviar dados ao grupo G, mesmo não sendo um membro deste grupo. Isto significa que uma estação maliciosa pode gerar dados e corromper uma aplicação *multicast*, bastando para isto conhecer o endereço do grupo *multicast* G. No serviço SSM, a ação maliciosa de uma estação só é possível se esta se passar pela fonte S sendo a técnica IP *spoofing*.

O mecanismo de contagem de pedidos de assinatura, que pode ser implementado por uma fonte SSM, possibilita a realização de eventos interativos. As emissoras de TV na Internet poderão realizar votações e enquetes sobre a sua programação e sobre o perfil dos seus usuários, enviando um número reduzido de pacotes. Esse mecanismo também poderá possibilitar a tarifação dos usuário do serviço de distribuição de TV.

A tarifação da TV na Internet por parte dos provedores de serviço da Internet se torna mais fácil. No SSM, é possível identificar a quem cobrar pelo serviço, pois existe a garantia de que apenas uma fonte transmite para um canal. Além disso, os roteadores dos provedores poderão implementar mecanismos de contagem de pedidos de assinatura a um dado canal específico. Os contratos entre os provedores de serviço e as emissoras de TV poderão ser baseados no número de assinantes de um canal.

Capítulo V

O Protótipo de Canal de TV e a Plataforma de Testes

Para analisar a viabilidade dos serviços IP Multicast e SSM, como mecanismos de distribuição de TV na Internet, foi projetada e construída uma plataforma de testes. Esta plataforma simula um domínio típico da Internet e possui um protótipo de emissora de TV.

Neste capítulo são apresentadas as ferramentas usadas na implementação do protótipo de canal de TV, da plataforma de testes e na realização dos experimentos envolvendo os dois serviços de distribuição *multicast*.

V.1 O Protótipo de Canal de TV

Um canal de TV na Internet deve ser composto, basicamente, por um servidor de vídeo, por uma pequena ilha de edição e por uma máquina de codificação.

No protótipo desenvolvido neste trabalho, um único computador é responsável por estas três funções. Este computador é uma máquina Dual Pentium IV Xeon 1.8 GHz rodando o sistema operacional Linux Red Hat, equipada com uma placa de captura de vídeo Brooktree bt848 e uma câmera de vídeo.

O protótipo roda dois programas servidores de vídeo: o Darwin Streaming Server e o VideoLAN. Este último é composto pelo VLC (VideoLAN Client) e pelo VLS (VideoLAN Server).

Através destes dois programas, o servidor é capaz de distribuir vídeo nos formatos MPEG-2, usado por todos os padrões de transmissão de TV digital terrestre e via satélite, e MPEG-4, desenvolvido para a transmissão de TV interativa na Internet.

V.1.1 Programas Servidores de Vídeo

Darwin Streaming Server

Um dos programas servidores de vídeo usados é o Darwin Streaming Server [24], da Apple. Este programa é um servidor de vídeo sob demanda para a plataforma Linux e só distribui vídeos codificados no formato MPEG-4.

O Darwin é capaz de distribuir arquivos armazenados na forma de fluxos (*streams*) de vídeo, criar listas de arquivos para exibição (*playlists*), fazer reflexões de transmissões ao vivo ou retransmissões de conteúdo de outros servidores. O servidor também é capaz de transmitir vídeo tanto em *unicast* quanto em *multicast*.

A requisição de conteúdo é feita pelos clientes através do protocolo RTSP (*Real Time Streaming Protocol*).

As listas de arquivos permitem que um ou mais arquivos sejam transmitidos de forma ininterrupta. Dessa forma, o servidor funciona como um canal de TV, no qual o cliente se conecta e começa a receber o conteúdo a partir do instante de tempo em que se encontra a transmissão e não no seu início. Para disponibilizar uma lista de arquivos, o servidor fornece aos clientes um arquivo de descrição da seção (SDP – *Session Description Protocol*), que contém as informações necessárias para obtenção do conteúdo.

A configuração do servidor Darwin é realizada via arquivos ou via navegador *web*. Somente a administração via navegador conta com a documentação de ajuda, que descreve todas as opções e recursos disponíveis no servidor. Com o navegador, também é possível observar estatísticas relativas ao funcionamento do servidor, como a porcentagem de ocupação da CPU, o número de conexões abertas e a quantidade de bytes servidos.

O servidor Darwin necessita de uma outra ferramenta para captura de vídeo ao vivo, uma vez que só é capaz de refletir transmissões ao vivo. A ferramenta mp4live [25] é uma opção. Sua função é criar um arquivo SDP com a informações sobre o vídeo capturado e direcionar o fluxo para o Darwin, que é capaz de refletir esse fluxo para os clientes.

VideoLAN

O servidor de vídeo também opera com a ferramenta VideoLAN [26], composta pelo VLC (VideoLAN Client) e pelo VLS (VideoLAN Server). Estes dois programas trabalham com o padrão MPEG (1, 2 e 4) e suportam a transmissão *multicast*. Além disso, estão disponíveis para diversas plataformas, incluindo Windows e Linux, e seus códigos fonte são abertos, o que permitirá a modificação dos aplicativos para que estes suportem o serviço SSM.

Tanto a ferramenta VLS quanto a VLC, inicialmente só um cliente, podem ser usadas para distribuir vídeo. O VLS não possui nenhuma interface gráfica. Sua configuração e feita via arquivos e o programa é executado por linha de comando. O VLS pode ser usado como um servidor de fluxos elementares MPEG-2 (DVD), canais de TV aberta, vídeos ao vivo, e fluxos originados por uma placa codificadora MPEG-2. Já o VLC possui uma interface gráfica amigável e pode distribuir vídeos nos formatos MPEG-2 e MPEG-4. As duas ferramentas podem usar os protocolos UDP e RTP para transmissão dos pacotes de vídeo e o RTSP para requisição de conteúdo.

A ferramenta VideoLAN também possui um serviço de informação de sessões, baseado nos protocolos de anúncio e de descrição de sessões, SAP/SDP. O servidor mini-SAP envia anúncios dos programas *multicast* ativos na rede e o cliente VLC os adiciona a sua lista de programas.

V.1.2 Programas Clientes de Vídeo

Todas as estações da plataforma de testes possuem programas clientes de vídeo compatíveis com os padrões MPEG-2 e MPEG-4 e com o protocolo RTSP.

Um dos programas clientes de vídeo usados é o mp4player, que é parte integrante do pacote MPEG4IP [25]. Este pacote inclui outras ferramentas úteis na criação, transmissão e recepção de vídeo em formato MPEG-4. O MPEG4IP é gratuito, de código aberto e funciona adequadamente nos sistemas operacionais Linux e Windows.

Outro cliente usado é o VLC [26], que recebe, decodifica e exibe fluxos de vídeo MPEG, padrões 1, 2 e 4. O VLC também é gratuito e de código aberto.

É importante ressaltar que nenhum desses clientes de vídeo possui suporte ao protocolo IGMPv3 e ao serviço SSM.

A biblioteca `live.com` [27], toda escrita em C++, permite o desenvolvimento de aplicações para distribuição de áudio e vídeo, usando os protocolos RTP/RTCP e RTSP e o padrão de codificação MPEG. Esta biblioteca possui suporte à versão 3 do protocolo IGMP e ao serviço SSM. Por isso, alguns dos programas de exemplo disponíveis foram usados para testar o SSM. Esse programas são servidores e clientes de vídeo. Entretanto, os clientes só acusam o recebimento dos pacotes e não exibem o conteúdo do vídeo.

V.1.3 Ferramentas de Videoconferência

O VIC (*Videoconference tool*) é uma ferramenta multi-plataforma para a realização de videoconferências na Internet. Inicialmente, foi desenvolvido para o uso no MBone (*Multicast Backbone*), mas também pode ser usado para transmissões ponto-a-ponto. Sua arquitetura é bastante flexível. As configurações da taxa de quadros por segundo, da qualidade da imagem transmitida, do codificador e da taxa de transmissão são facilmente ajustadas.

É importante ressaltar que o VIC é o responsável apenas pela transmissão e recepção do vídeo em uma conferência multimídia. O áudio, o quadro branco compartilhado e o controle de sessão são responsabilidades, respectivamente, do RAT (*Robust Audio Tool*), do WB (*Whiteboard Tool*) e do SDR (*Session Directory Tool*).

Existem várias versões modificadas do VIC. A usada neste trabalho foi a versão do projeto OpenMash [28], que possui suporte ao serviço SSM. Esta versão ainda não foi completamente testada, por isso só está disponível no repositório de arquivos do projeto. A sintaxe para execução do programa foi modificada. Na versão original, os parâmetros necessários para o funcionamento do VIC eram o endereço, *multicast* ou *unicast*, do destinatário e a porta de comunicação (ex:`./vic 224.1.1.1/10000`). Na versão modificada está sintaxe também é válida. Entretanto, para usar o serviço SSM, o endereço da fonte deve ser informado (ex:`./vic 192.168.10.3@232.1.1.1/10000`).

V.2 A Plataforma de Testes

A plataforma de testes foi construída usando a estrutura física disponível no laboratório do Grupo de Teleinformática e Automação (GTA). Cada computador usado na plataforma está conectado ao comutador Ethernet do laboratório. Todos os enlaces utilizados são Ethernet, com capacidades de 10 ou 100 Mbps.

Todas as máquinas da plataforma, inclusive os roteadores, são computadores pessoais que utilizam o sistema operacional Linux Red Hat, com a versão 2.4.20 do núcleo (*kernel*).

Os computadores não são usados exclusivamente na plataforma, pois são estações de trabalho dos alunos do GTA. Como estes computadores deveriam continuar conectados à rede de trabalho do laboratório, foram instaladas interfaces de rede adicionais em alguns deles. Cada interface de rede dos computadores usados na plataforma recebe um endereço da faixa `192.168/16` reservada para experimentos locais. A distribuição dos endereços foi feita de acordo com a sub-rede onde a máquina estava localizada.

V.2.1 Topologia

A plataforma de testes simula uma rede *backbone* contendo elementos típicos de um domínio na Internet. Sua topologia é mostrada na Figura V.1.

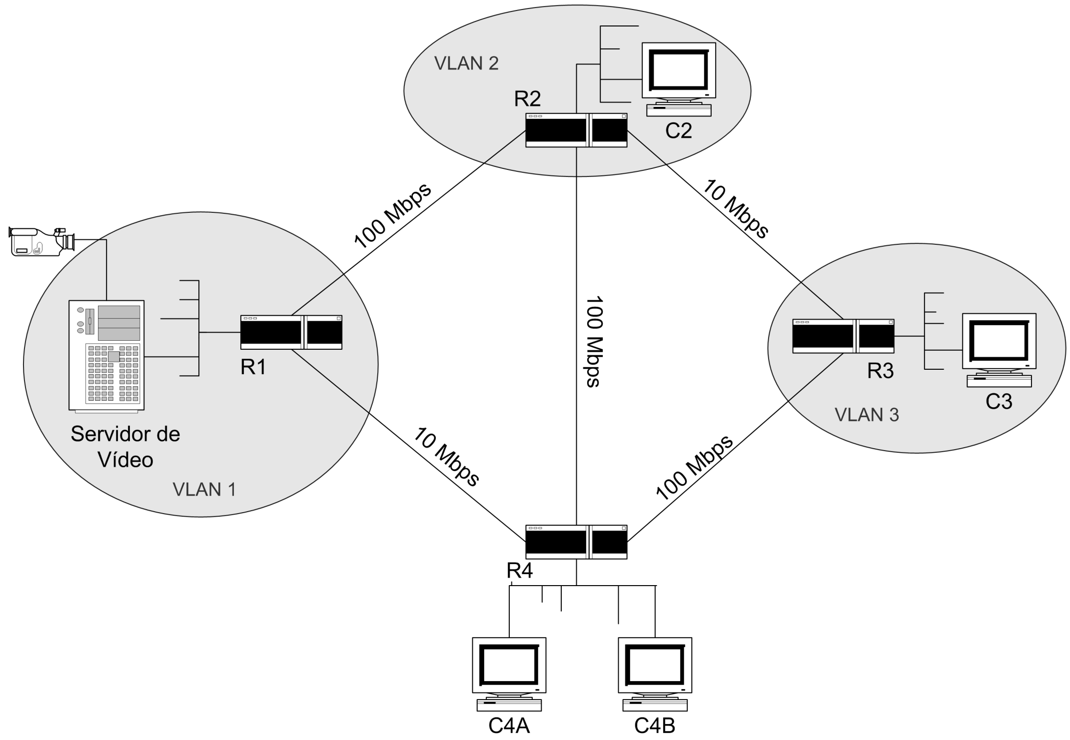


Figura V.1: A topologia da plataforma de testes.

A topologia da plataforma de testes foi projetada para permitir a existência de múltiplos caminhos entre a fonte e os receptores. Dessa forma, foi possível observar algumas situações críticas na construção da árvore de distribuição e no funcionamento dos protocolos de roteamento, que serviram de parâmetro para analisar o desempenho desses protocolos.

A plataforma é composta por quatro roteadores, por um servidor de vídeo e por clientes. O servidor S está conectado ao roteador R_1 por meio de uma VLAN (*Virtual Local Area Network*). O uso das VLANs é necessário para isolar o tráfego da plataforma de testes da rede de trabalho do laboratório e para criar enlaces ponto-a-ponto entre algumas máquinas da plataforma. Existe pelo menos um cliente conectado aos roteadores

R_2, R_3 e R_4.

O protocolo IGMP permite que apenas um roteador local atue como "interrogador". De acordo com o mecanismo implementado pelo protocolo, se mais de um roteador envia mensagens de interrogação na rede, estas só devem ser respondidas ao roteador com o menor endereço IP. Como a interface que interconecta os roteadores R_2, R_3 e R_4 e seus clientes é a mesma que os conecta à rede de trabalho do laboratório, eles pertencem a mesma rede local. Portanto só um deles receberia respostas às suas interrogações. Para solucionar este problema foram criadas duas VLANs que interconectam, respectivamente, o roteador R_2 ao cliente C_2 e o roteador R_3 ao cliente C_3. Somente o roteador R_4 permite o acesso dos clientes C_{4A} e C_{4B} localizados na rede de trabalho do laboratório ao servidor de vídeo.

Uma vez que a maior parte do tráfego da plataforma flui do servidor de vídeo para os clientes, os enlaces de 10 Mbps entre os roteadores R_1 e R_4 e os roteadores R_2 e R_3 representam os gargalos da plataforma de testes.

O roteamento na plataforma de testes é dinâmico. Todos os roteadores rodam o GateD, um programa de roteamento executado em *background* pelo sistema operacional. Todos os roteadores estão configurados com o protocolo de roteamento *unicast* OSPF (*Open Shortest Path First*) [17]. O protocolo de roteamento *multicast* usado é determinado pelo teste a ser realizado. Os protocolos que podem ser usados são o DVMRP, o PIM-DM, o PIM-SM e o PIM-SSM. Mais detalhes sobre a ferramenta GateD serão apresentados na Seção V.2.3.

Para que os clientes possam enviar e receber corretamente os pacotes, é necessário criar uma entrada na tabela de roteamento de cada um deles. Esta entrada deve indicar o roteador a ser usado como portal para a plataforma de testes e a interface de saída para este roteador. Isto pode ser feito usando a ferramenta ip do Linux, através do comando "`ip route add 192.168.0.0/16 via` *portal* `dev` *interface*". Para cada cliente, os parâmetros `portal` e `interface` correspondem, respectivamente, ao endereço IP do roteador ao qual o cliente está conectado e a interface de rede que o conecta a este roteador.

Também é necessária a adição das rotas para as faixas de endereço IP Multicast e SSM, pois só dessa forma as aplicações conseguem associar um grupo *multicast* à interface desejada. Isto também pode ser feito com a ferramenta ip, através dos comandos "`ip route add 224.0.0.0/4 via` *portal* `dev` *interface*" e "`ip route add 232.0.0.0/8 via` *portal* `dev` *interface*". Estes comandos devem ser adicionados ao arquivo `/etc/rc.local` para que sejam executados automaticamente, quando o sistema operacional for reiniciado.

V.2.2 Suporte ao Serviço SSM no Linux

O serviço SSM atualmente é implementado pelos protocolos IGMPv3 e PIM-SSM. Todos os receptores e os roteadores de primeiro salto devem implementar a versão 3 do protocolo IGMP. O funcionamento deste protocolo pode ser dividido em "lado estação" e "lado roteador".

O lado roteador é implementado pela ferramenta GateD. Já o lado estação deve ser implementado pelo sistema operacional. Na versão 2.4.20 do núcleo do sistema operacional Linux, usada nos computadores da plataforma de testes, o suporte ao IGMPv3 não é oferecido. Portanto foi necessário aplicar um remendo (*patch*) ao código fonte do núcleo do sistema operacional. Este remendo foi desenvolvido por Vinay Kulkarni e encontra-se disponível em [29]. Após a aplicação do remendo, não é necessário habilitar nenhuma opção de configuração no núcleo do sistema operacional. Entretanto o núcleo do sistema deve ser recompilado.

O suporte ao roteamento *multicast* e ao protocolo PIM já são oferecidos pela versão 2.4.20 do Linux. Para isto, em todos os roteadores as opções de configuração do núcleo `CONFIG_IP_MROUTE,` `CONFIG_IP_PIMSM_V1` e `CONFIG_IP_PIMSM_V2` devem estar habilitadas. Com estas opções habilitadas, o núcleo do sistema deve ser recompilado.

A opção `CONFIG_IP_MULTICAST` vem habilitada por padrão na distribuição Red Hat do Linux. Ele deve estar habilitada em todas as estações para que o serviço de rede saiba como tratar os pacotes *multicast*.

É importante ressaltar que o suporte dos protocolos pelo núcleo é necessário para que o sistema operacional saiba como tratar os pacotes deste protocolos. A responsabilidade pela implementação dos protocolos é da ferramenta de roteamento.

V.2.3 Roteamento

O roteamento na plataforma de testes é dinâmico. Os roteadores aprendem as rotas para cada sub-rede através de informações trocadas entre si. Todos os roteadores da plataforma de testes são computadores pessoais rodando o sistema operacional Linux e a ferramenta de roteamento GateD.

A Ferramenta GateD

Atualmente, o GateD é uma ferramenta de roteamento comercial e de código fonte fechado. No entanto, a versão usada neste projeto é gratuita e foi obtida através do sítio do fabricante na Internet (`http://www.nexthop.com`).

Esta ferramenta foi escolhida por ser a única a implementar os protocolos IGMPv3 e PIM-SSM, necessários para a implantação do serviço SSM.

O GateD também implementa os protocolos de roteamento *unicast* RIP, OSPF e BGP e ainda os protocolos de roteamento *multicast* DVMRP e os modos denso e esparso do PIM. No caso do PIM, o GateD requer o uso do protocolo OSPF.

A configuração do GateD pode ser feita através de um arquivo XML (*eXtensible Markup Language*) ou através de uma interface de linha de comando, semelhante à dos roteadores Cisco. Todas as configurações feitas através da interface de linha de comando são salvas no arquivo `/etc/gated.conf`. Para acessar a interface de linha de comando, basta executar o comando `cligated`, com o GateD rodando em segundo plano. Existem dois modos de configuração. O modo usuário só permite a visualização das tabelas de roteamento e de algumas estatísticas referentes a cada protocolo. Já o modo privilegiado permite a configuração de todos os parâmetros do GateD. Para execução do modo privilegiado é necessário que o usuário entre com una senha, que é registrada durante a instalação da ferramenta. Todos os parâmetros de configuração de cada protocolo implementado com o GateD podem ser encontrados no manual, distribuído junto com o programa [30].

Para que o PIM possa usar as rotas obtidas pelo OSPF, dentro da configuração do roteador OSPF é necessário ativar a opção `multicast-rib`. Dessa forma, a base de dados do OSPF pode ser acessada pelo PIM para a construção da árvore de distribuição.

O GateD também possui um programa de gerenciamento, o gdc. Com este programa é possível iniciar, parar, reiniciar e verificar se o GateD está sendo executado.

V.2.4 Geração de Tráfego e Medição

Para que fosse possível realizar testes de desempenho entre os serviços de distribuição *multicast* e entre os protocolos de roteamento, foram usadas ferramentas que simulam latência, vazão e perdas estatísticas. Todas as ferramentas usadas na geração e medição de tráfego rodam no Linux, são livres e de código aberto.

A Ferramenta iperf

O iperf [31] é uma ferramenta cliente-servidor usada para medir o desempenho da rede. Com esta ferramenta é possível, determinar largura de banda disponível, a variação do atraso e a perda de pacotes. Neste trabalho, o iperf foi usado para simular o tráfego de vídeo, transmitindo tanto pacotes *unicast* quanto pacotes *multicast*. Foram enviados pacotes UDP, a uma taxa de bits constante.

O cliente iperf é o responsável pela geração do tráfego. Pela linha de comando, são configurados parâmetros como o endereço do cliente *unicast*, ou do grupo *multicast*, o tipo de protocolo de transporte, a banda passante e o valor do campo TTL. No caso da transmissão *multicast*, o valor do campo TTL é importante, já que, por padrão nesta ferramenta, ele vale 1 para todo pacote.

O servidor iperf é o responsável pela recepção do tráfego gerada pelos clientes. Pela linha de comando, é possível configurar o tipo de protocolo de transporte usado e, no caso da transmissão *multicast*, o endereço do grupo. O servidor é o responsável por calcular e exibir a banda passante, a taxa de perda e a variação do atraso.

O iperf usa a versão 2 do protocolo IGMP para sinalizar ao roteador local o interesse na recepção de dados de um grupo *multicast*. Como a ferramenta não possui suporte ao IGMPv3, o servidor iperf não pode ser usado para medir a variação do atraso na árvore de distribuição do PIM-SSM. Para o cliente, o IGMPv3 não é importante, uma vez que a estação pode enviar dados para qualquer endereço de grupo. No SSM, cabe ao receptor a aceitação ou não do tráfego enviado a um endereço *multicast*.

As Ferramentas ssmsend e ssmrec

Os programas ssmsend e ssmrec [32] são usados para testar o funcionamento do serviço SSM, no Linux.

A ferramenta ssmsend emite pacotes *multicast* em intervalos regulares de tempo e imprime, opcionalmente, o número de seqüência destes pacotes na tela. Esta ferramenta não requer o suporte ao IGMPv3.

A ferramenta ssmrec permite que uma estação se torne assinante de um canal *multicast* específico. Para cada pacote recebido, a ferramenta imprime o endereço IP da fonte, o número de seqüência do pacote e o atraso no caminho entre a fonte e o assinante (*one-way delay*). Esta ferramenta requer o suporte do núcleo do sistema operacional ao protocolo IGMPv3.

As Ferramentas ntop, tcpdump e ethereal

As ferramentas ntop, tcpdump e ethereal foram usadas para monitorar e medir a taxa de utilização de cada enlace durante a realização dos testes de desempenho. Dessa forma, foi possível comparar a transmissão ponto-a-ponto e a transmissão *multicast* e analisar a construção da árvore de distribuição dos protocolos de roteamento *multicast*.

O ntop é um programa usado para gerenciar a rede. Com ele é possível classificar o tráfego em um determinado enlace, obter a matriz de tráfego entre diferentes origens e destinos e observar a taxa de utilização de um enlace, a cada minuto e a cada hora. O acesso ao programa é feito através de um navegador *web*.

O tcpdump e o ethereal são programas farejadores (*sniffers*). Um farejador monitora e registra o tráfego de dados em uma interface específica. Para analisar todo o tráfego que passa por um segmento de rede, mesmo os pacotes que não são destinados aquela máquina, a interface de rede deve operar em modo promíscuo. Os farejadores são úteis para a detecção e solução de problemas na rede. Entretanto, podem ser usados para capturar senhas e dados confidencias, que trafegam pela rede sem qualquer proteção. O tcpdump é um programa de linha de comando. Já o ethereal possui uma interface gráfica amigável e de fácil utilização.

Capítulo VI

Testes e Resultados

A plataforma construída neste trabalho permitiu a realização de testes comparativos entre a transmissão de dados ponto-a-ponto e a transmissão *multicast*. Os testes também envolveram a comparação entre os serviços IP Multicast e SSM para distribuição de TV na Internet. Foi feita ainda uma análise do desempenho dos principais protocolos de roteamento *multicast*.

Neste capítulo, a descrição e os resultados dos testes realizados são apresentados e discutidos.

VI.1 Eficiência da Distribuição Multicast

Uma das vantagens da distribuição *multicast* é a redução da taxa de utilização da rede, uma vez que não é necessário enviar um pacote para cada estação interessada na informação de uma fonte de tráfego. O uso de um serviço *multicast* eficiente, em termos de banda passante, é essencial para a implantação de um serviço de TV na Internet, pois devido ao grande número de usuários envolvidos, o uso de múltiplas conversações ponto-a-ponto é inviável.

Para comprovar a eficiência da distribuição *multicast*, foi medida a taxa média de ocupação de um enlace da plataforma de testes em função da variação do número de clientes que requisitam o conteúdo de uma fonte.

Em um primeiro momento, para atender os clientes, são usadas transmissões ponto-a-ponto entre cada cliente e a fonte S. Em seguida, cada cliente que deseja receber o tráfego vindo da fonte deve se inscrever em um grupo *multicast*.

O enlace escolhido para o teste conecta a fonte S ao roteador R_1, como mostra a Figura VI.1. Este enlace foi escolhido, pois todo o tráfego gerado pela fonte na direção dos clientes passa pelo roteador R_1.

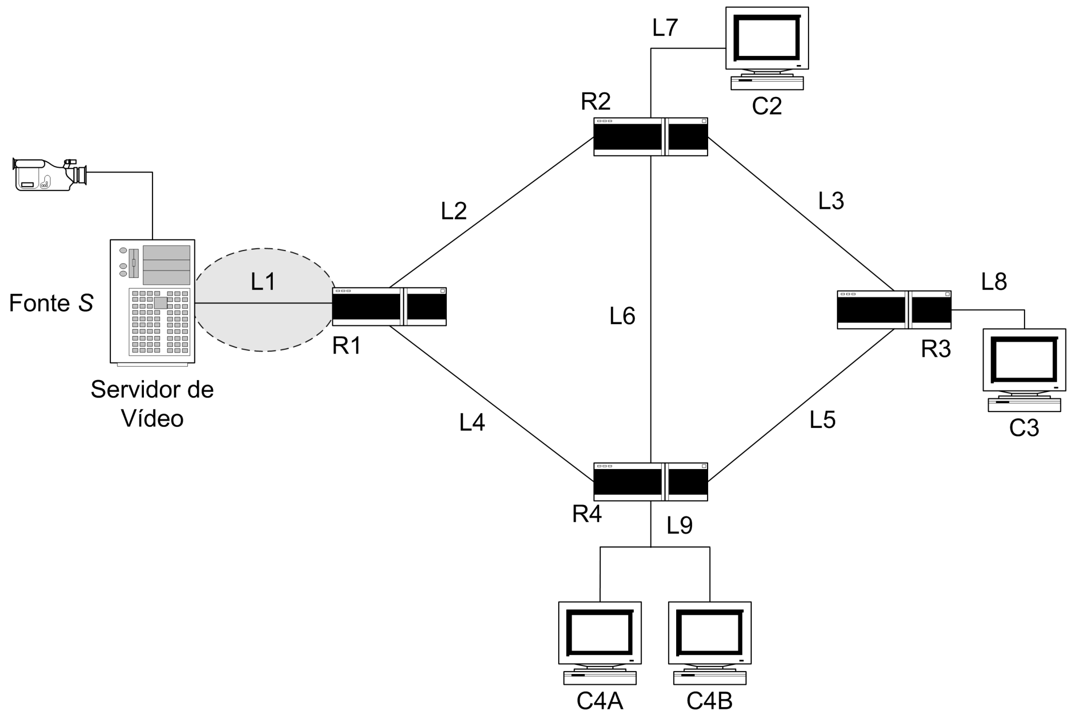


Figura VI.1: A topologia da plataforma de testes - enlace L_1.

Para cada cliente, durante um tempo de t segundos, foram transmitidos pacotes UDP a uma taxa de 4,0 Mbps. Este valor corresponde à taxa de transmissão média de um vídeo codificado em MPEG-2, com resolução de 720×480 *pixels* e taxa de 30 quadros por segundo. Estes valores são referentes ao padrão NTSC (*National Television System Committee*) [33]. O tráfego gerado pelo MPEG possui taxa de bits variável. Entretanto o uso de um tráfego de taxa de bits constante, gerado pela ferramenta iperf, não interfere nos resultados deste trabalho.

Os protocolos usados neste teste foram o OSPF, para o roteamento *unicast*, e o PIM-DM, para o roteamento *multicast*. O protocolo de roteamento não influi nos resultados deste teste, uma vez que só um enlace está sendo monitorado e não está se medindo a taxa de perda ou o atraso dos pacotes enviados.

O experimento foi realizado da seguinte forma. Primeiro somente um cliente requisita o conteúdo da fonte. Em seguida dois clientes requisitam o conteúdo de S e assim sucessivamente. Para cada situação, o mesmo teste foi realizado cinco vezes. A Figura VI.2 mostra as curvas para a taxa de utilização do enlace L_1, em função do aumento do número de clientes que requisitam o conteúdo da fonte S. Cada ponto do gráfico, corresponde à média das cinco rodadas de testes.

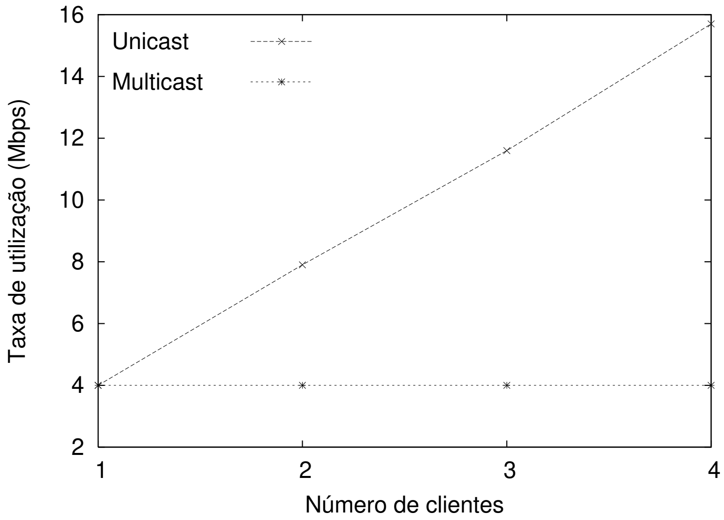


Figura VI.2: A taxa de utilização da rede em função do número de clientes.

Pela Figura VI.2, é possível observar que à medida que o número de clientes aumenta, a taxa de utilização da rede para a distribuição *multicast* se mantém constante, enquanto que para a transmissão ponto-a-ponto ela cresce linearmente até atingir a capacidade máxima do enlace. Isto comprova a eficiência do serviço *multicast* e justifica o seu uso para a implantação de um serviço de distribuição de TV na Internet.

VI.2 O Serviço SSM

O serviço SSM é especializado na distribuição de aplicações de fonte única. Seu modelo de serviço reduz a conversação *multicast* para um-para-muitos e usa o conceito de canal como abstração de serviço. Isto torna o problema da alocação de endereços local à fonte, permite o controle de acesso ao canal e garante a integridade das aplicações, exigências para a distribuição de TV na Internet não atendidas pelo serviço IP Multicast.

Para observar as características do modelo de serviço SSM e compará-lo com o IP Multicast, foram realizados testes envolvendo múltiplas fontes transmitindo para o mesmo endereço de grupo e diversos clientes, que requisitam o conteúdo destas fontes.

O objetivo dos experimentos é simular uma transmissão de TV, em que somente alguns usuários têm direito a assinar o canal e, portanto, conhecem o endereço S da fonte. O endereço de destino do canal G é divulgado publicamente pela emissora de TV, o que possibilita a tentativa de uma ação maliciosa por parte de um usuário mal-intencionado.

Tanto a fonte, quanto os clientes, usam uma versão modificada da ferramenta VIC [28]. O usuário mal-intencionado gera tráfego com a ferramenta iperf [31].

Nos testes, o servidor de vídeo, fonte S, transmite conteúdo de vídeo, para um endereço de destino G, exclusivo da faixa SSM. Um cliente C_1 que tem direito a assinar o canal, envia à fonte uma mensagem de assinatura (S,G) e passa a receber o vídeo. Um outro cliente C_2, ao qual não é permitido o acesso ao canal e, portanto, não conhece o endereço S da fonte, tenta receber o conteúdo de vídeo apenas enviando à fonte uma mensagem de inscrição ao grupo G. Não houve sucesso nesta tentativa. O pedido de inscrição ao grupo G não foi encaminhado à fonte pelo próprio roteador local do cliente C_2, uma vez que pertence à faixa exclusiva SSM.

Em outra situação um usuário mal-intencionado M tenta interferir na transmissão de vídeo. Para isso, ele atua como uma fonte enviando dados ao endereço de destino G, usado pelo servidor de vídeo. Esta outra tentativa também não obteve sucesso. Como o interesse do cliente C_1 é no canal (S,G) e os roteadores PIM-SSM só constroem árvores de distribuição por fonte, os pacotes enviados por M ao endereço G não foram encaminhados pelo seu próprio roteador local, uma vez que não havia nenhum receptor interessado no conteúdo da “fonte” M.

Estes mesmos testes foram realizados com o servidor de vídeo enviando o conteúdo para um endereço de grupo Y, da faixa IP Multicast. Nas duas situações descritas anteriormente, as ações maliciosas foram bem sucedidas. Um cliente sem autorização consegue receber o conteúdo enviado pelo servidor e um usuário mal-intencionado interfere na transmissão do vídeo.

É importante ressaltar que o modelo de serviço SSM facilita o controle de acesso ao grupo, mas não o garante. É possível que usuários mal-intencionados obtenham o endereço da fonte, farejando os enlaces por onde trafegam os pacotes endereçados ao destino conhecido G. Da mesma forma, um usuário conhecendo o endereço S, pode se passar pela fonte, usando a técnica IP *spoofing*, e enviar dados ao canal (S,G).

VI.3 Protocolos de Roteamento Multicast

Um protocolo de roteamento *multicast* é o responsável pela criação da árvore de distribuição do tráfego *multicast*. Cada um dos protocolos constrói um tipo de árvore usando diferentes mecanismos. De acordo com o tipo de árvore e com o tipo de mecanismo usado na construção desta, é possível determinar qual o protocolo mais adequado para atender uma determinada aplicação.

Os testes realizados simulam a transmissão de um programa de TV na Internet. Enquanto a fonte distribui o vídeo e os enlaces da rede são monitorados, os clientes se inscrevem e deixam um grupo. Dessa forma, é possível observar os mecanismos de construção e adaptação da árvore de distribuição de um protocolo de roteamento *multicast*. A configuração da plataforma usada nos testes comparativos entre os protocolos de roteamento é mostrada na Figura VI.3.

O cenário usado para os testes é descrito a seguir:

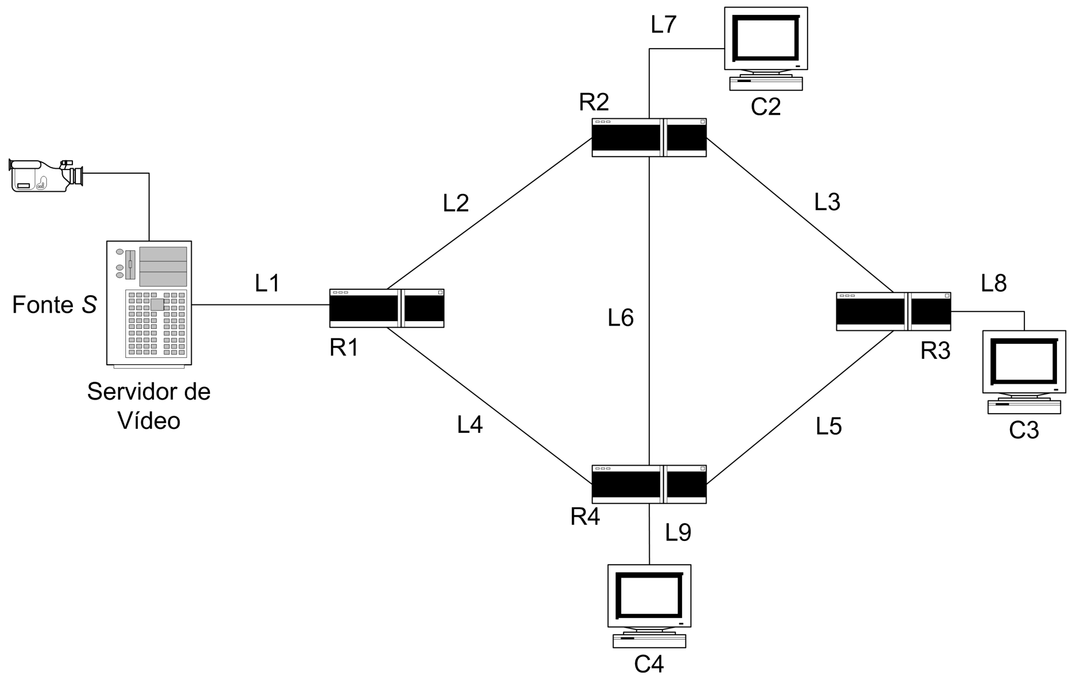


Figura VI.3: A configuração da plataforma.

- em $t = 0$ s, a fonte S começa a distribuir os pacotes de vídeo;
- em $t = 20$ s, o cliente C_2 requisita a inscrição em um grupo, ou a assinatura de um canal *multicast*;
- em $t = 40$ s, outro cliente, C_3, requisita a inscrição no mesmo grupo, ou a assinatura do mesmo canal *multicast* de C_2;
- em $t = 60$ s, um terceiro cliente, C_4, requisita a inscrição no mesmo grupo, ou a assinatura do mesmo canal *multicast* de C_2 e C_3;
- entre $t = 60$ s e $t = 120$ s, não há nenhuma modificação na árvore de distribuição;
- em $t = 120$ s, o cliente C_3 deixa o grupo ou cancela a assinatura do canal *multicast*;
- em $t = 160$ s, o cliente C_2 deixa o grupo ou cancela a assinatura do canal *multicast*;
- em $t = 180$ s, a fonte S deixa de distribuir o vídeo.

Para simular este cenário, foram usados *scripts* que executam as ferramentas de medição, geração e recepção de tráfego nos roteadores, clientes e no servidor de vídeo. Todas as máquinas da plataforma de testes estão com os relógios sincronizados, através do protocolo NTP (*Network Time Protocol*).

Tabela VI.1: Parâmetros de configuração dos protocolos usados nos testes.

Protocolos	Parâmetros	Valor
IGMP	`query-interval`	125 s
	`query-max-response-interval`	125 s
OSPF	`hello-interval`	10 s
	`dead-interval`	40 s
	`poll-interval`	120 s
	`multicast-rib`	ativada
DVMRP	`prune-lifetime`	7200 s
PIM	`hello-interval`	30 s
	`hello-holdtime`	105 s
	`assert-holdtime`	180 s
	`join-prune-interval`	60 s
	`join-prune-holdtime`	210 s
PIM-DM	`graft-retry`	3 s
	`source-lifetime`	210 s
	`state-refresh-interval`	60 s
PIM-SM	`threshold`	1000 bps
	`threshold-dr`	1000 bps
	`threshold-rp`	1000 bps
	`probe-interval`	5 s
	`rp-candidate-holdtime`	150 s

Este cenário foi usado para avaliar o desempenho dos protocolos de roteamento *multicast* DVMRP, PIM-DM, PIM-SM e PIM-SSM. O protocolo PIM, precisa de um protocolo de roteamento *unicast* para construir a sua árvore de distribuição. Nos testes, por limitação da ferramenta GateD, foi utilizado o protocolo de estado de enlace OSPF. O roteador R_2 foi escolhido como ponto de encontro em todos os experimentos envolvendo o PIM-SM.

Os valores usados para os temporizadores e outros parâmetros de alguns destes protocolos são os valores padrão da ferramenta GateD. Alguns destes valores são apresentados na Tabela VI.1. A alteração de alguns destes valores pode modificar o comportamento de cada protocolo.

Como no teste anterior, a fonte S transmite pacotes UDP a uma taxa de 4,0 Mbps, simulando um vídeo codificado em MPEG-2 com esta taxa média. O tráfego UDP foi gerado através da ferramenta iperf. Os clientes usam o iperf, no caso dos protocolos IP Multicast, e o ssmrec, no caso do SSM. Para monitorar a rede é usada a ferramenta tcpdump, que permite a gravação em arquivo dos pacotes que trafegam por cada enlace da plataforma de testes.

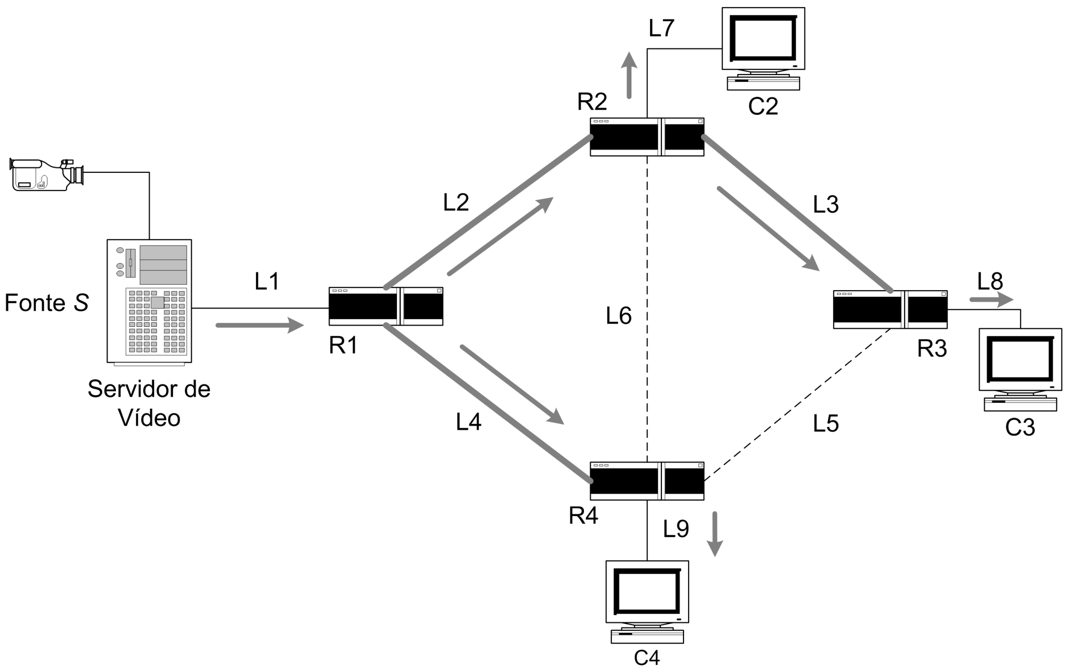


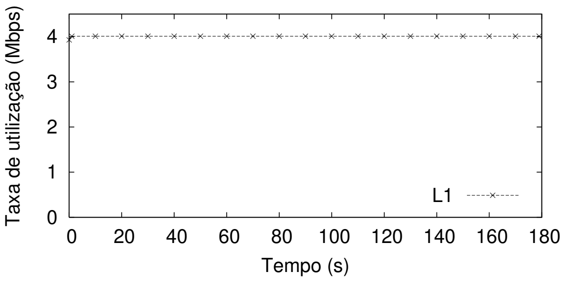
Figura VI.4: A árvore de distribuição construída.

A árvore de distribuição construída por todos os protocolos durante os testes é a mesma, mostrada na Figura VI.4, apesar do roteador R_3 possuir dois caminhos com a mesma distância para a fonte, ou seja, dois saltos. No caso do DVMRP, que constrói a sua própria tabela de roteamento *unicast*, R_3 escolhe R_2 como próximo salto, pois o roteador R_2 possui o menor endereço IP, se comparado ao endereço de R_4. No caso do OSPF, que é usado pelo PIM para construir a árvore de distribuição, o protocolo foi configurado para usar o número de saltos como métrica, mas não para usar múltiplos caminhos até um dado destino. Dessa forma, R_2 é escolhido como próximo salto de R_3 para chegar à fonte, por possuir o menor endereço IP.

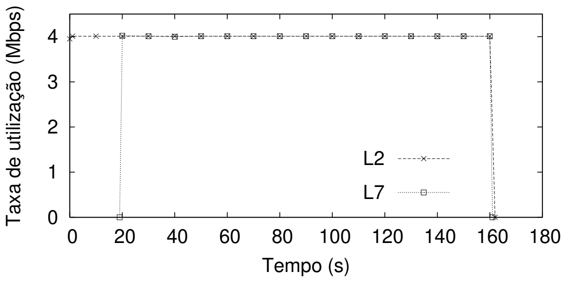
O cenário usado na simulação do programa de TV permitiu a observação do comportamento de cada um dos protocolos durante a construção e adaptação da árvore de distribuição. As Figuras de VI.5(a) a VI.8(d) mostram a ocupação de todos os enlaces da plataforma de testes durante a simulação da transmissão de TV, para um determinado protocolo. Para cada protocolo, este mesmo experimento foi realizado cinco vezes. Os valores apresentados da Figura VI.5(a) até a Figura VI.8(d) correspondem à média dos resultados destes cinco experimentos.

As Figuras VI.5(a), VI.6(a), VI.7(a) e VI.7(a) mostram o comportamento do enlace L_1, respectivamente, para os protocolos DVMRP, PIM-DM, PIM-SM e PIM-SSM. É possível notar que este enlace está sempre ocupado, pois é por ele que flui todo o tráfego do servidor de vídeo em direção aos clientes. Outro detalhe, os enlaces L_5 e L_6 só foram ocupados com pacotes de dados pelo protocolo PIM-DM, como será visto adiante. Para os demais protocolos, como não existe tráfego, os gráficos de ocupação destes enlaces foram omitidos.

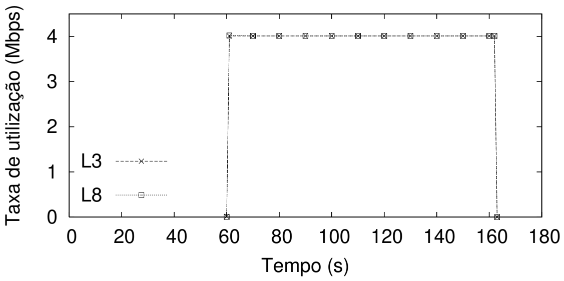
Quando estão dispostos dois a dois nos gráficos, Figura VI.5(a) a Figura VI.8(d), os enlaces no mesmo gráfico possuem comportamento semelhante. O enlace L_2 deve estar ocupado a partir do instante $t = 20$ s, pois ele faz parte dos ramos da árvore que ligam a fonte S aos clientes C_2, primeiro a solicitar o conteúdo de S, e C_3, último cliente a deixar o grupo. Já o enlace L_7 só deve estar ocupado enquanto o cliente C_2 tem interesse nos pacotes vindos de S. Os enlaces L_3 e L_8 só devem estar ocupados quando o cliente C_3



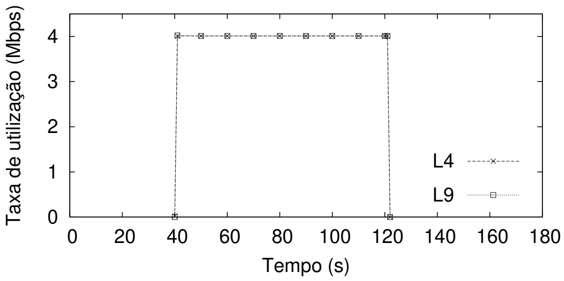
(a) Enlace L_1.



(b) Enlaces L_2 e L_7.



(c) Enlaces L_3 e L_8.



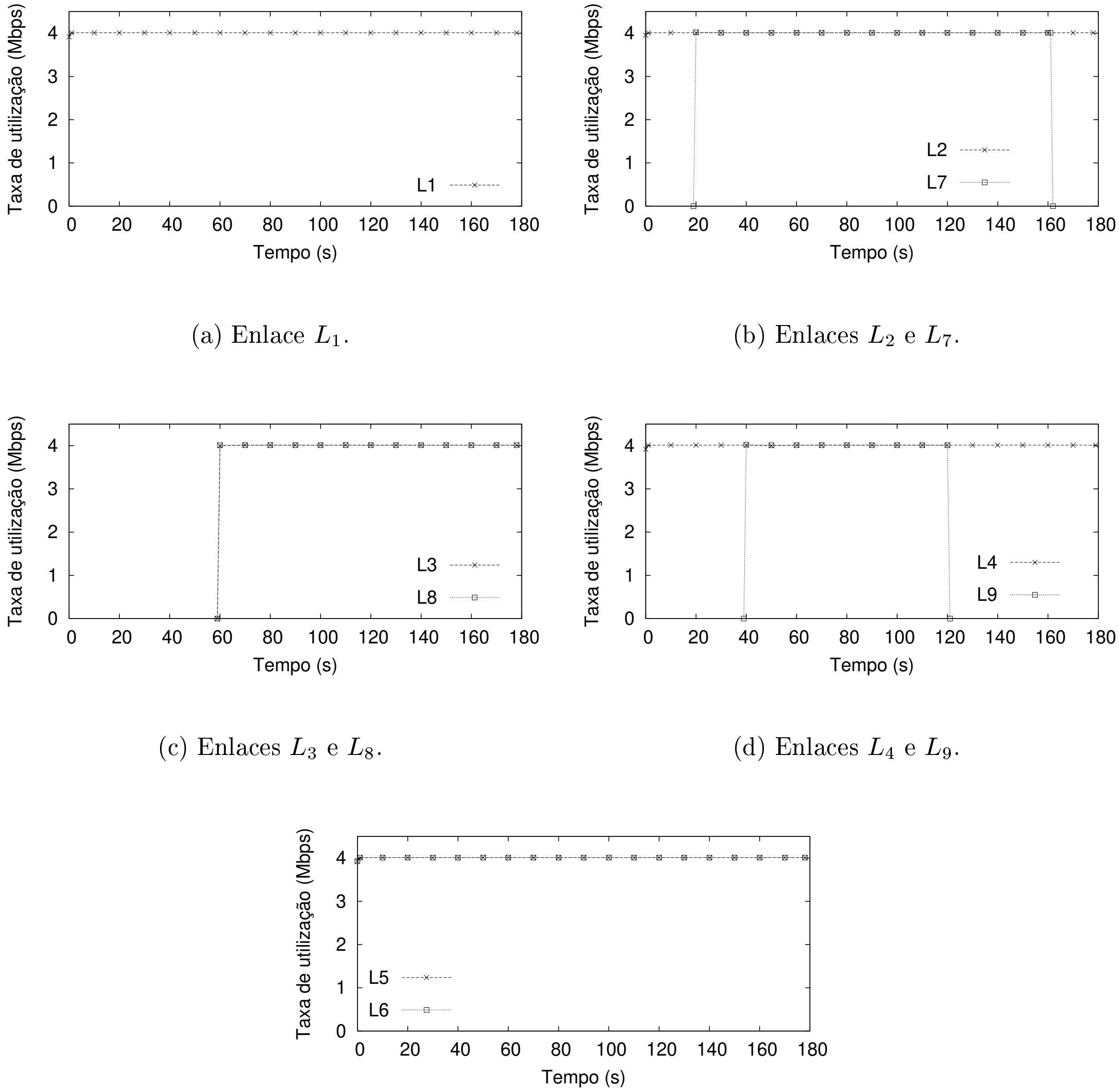
(d) Enlaces L_4 e L_9.

Figura VI.5: Ocupação dos enlaces para o DVMRP.

requisita o conteúdo da fonte S. Da mesma forma, os enlaces L_4 e L_9 só devem estar ocupados quando o cliente C_4 tem interesse no tráfego da fonte S.

Ao observar as Figuras de VI.5(a) a VI.8(d), nota-se que os protocolos de roteamento que usam um mecanismo de inundação-e-poda para construir a árvore de distribuição, neste caso o DVMRP e o PIM-DM, possuem uma maior taxa de utilização da rede.

O DVMRP inunda o enlace L_2, mostrado na Figura VI.5(b), durante os 20 s iniciais do experimento, quando não existe nenhum cliente interessado no conteúdo de S. Além disso, como mostra a Figura VI.5(c), o cliente C_3, em $t = 160$ s, deixa de receber pacotes *multicast* vindos da fonte, mesmo ainda sendo um membro do grupo. Isto ocorre, pois neste mesmo instante o cliente C_2 envia uma mensagem ao seu roteador local, no caso R_2, para deixar o grupo. Mesmo sabendo que R_3 o usa como pai na árvore de distribuição, o roteador R_2 envia uma mensagem de poda na direção da fonte pelo enlace L_2. Este



(a) Enlace L_1.

(b) Enlaces L_2 e L_7.

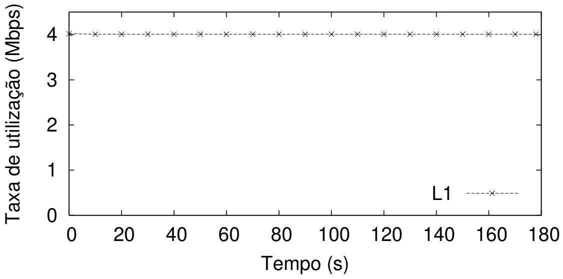
(c) Enlaces L_3 e L_8.

(d) Enlaces L_4 e L_9.

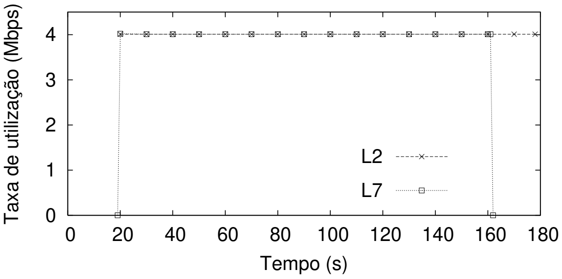
(e) Enlaces L_5 e L_6.

Figura VI.6: Ocupação dos enlaces para o PIM-DM.

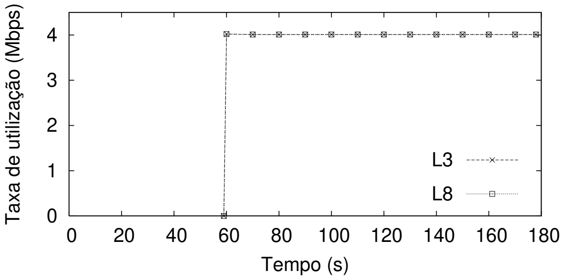
enlace, então, é podado da árvore de distribuição. Com isso, o cliente C_3 deixa de receber o tráfego de S, uma vez que o enlace L_2 faz parte do ramo da árvore entre a fonte e este cliente. Isto mostra que o protocolo DVMRP não funcionou corretamente durante a simulação da transmissão de TV, o que indica uma falha na implementação da ferramenta GateD.



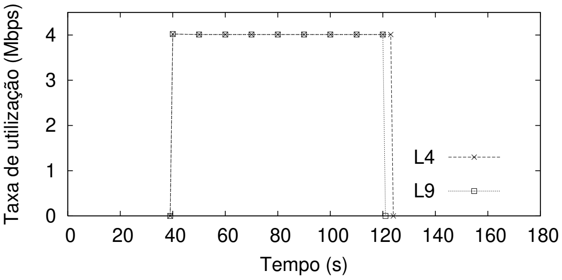
(a) Enlace L_1.



(b) Enlaces L_2 e L_7.



(c) Enlaces L_3 e L_8.

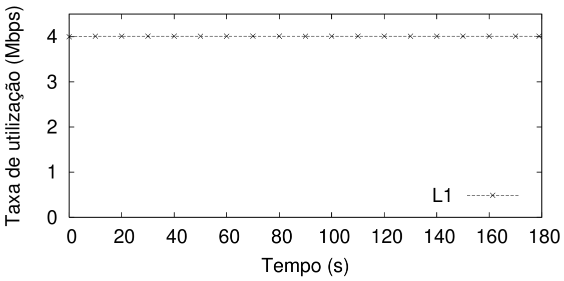


(d) Enlaces L_4 e L_9.

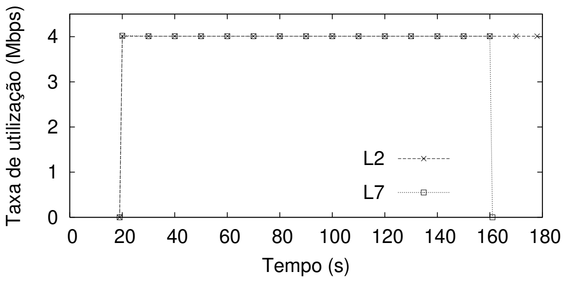
Figura VI.7: Ocupação dos enlaces para o PIM-SM.

O outro protocolo que usa o mecanismo de inundação-e-poda para construir a árvore de distribuição, o PIM-DM, apresenta a maior taxa de utilização da rede. Durante todo o experimento os enlaces L_5 e L_6 são inundados desnecessariamente, como mostra a Figura VI.6(e). O mesmo ocorre para os enlaces L_2 e L_4, durante os 20 s iniciais e também para o enlace L_4, entre $t = 120$ s e $t = 180$ s. A ocupação desnecessária dos enlaces ocorre, pois os roteadores R_2 e R_4 deixam de enviar mensagens de poda nas suas interfaces que não foram indicadas como a "melhor" no teste RPF. Apesar de manter a rede ocupada desnecessariamente, se comparado ao DVMRP, o PIM-DM funcionou corretamente, uma vez que todos os clientes recebem o tráfego da fonte, enquanto este é solicitado.

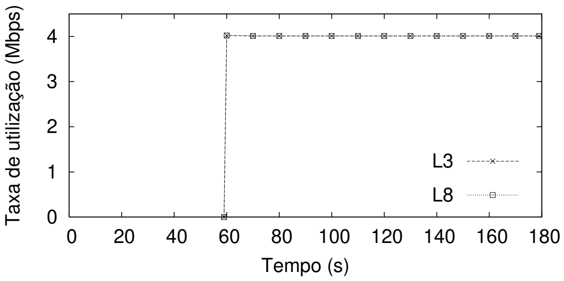
Os dois protocolos que usam o mecanismo de inscrição e poda explícitas para a construção da árvore de distribuição, o PIM-SM e o PIM-SSM, apresentam as menores taxas de utilização da rede, durante todo o experimento. No caso do PIM-SM, foi construída



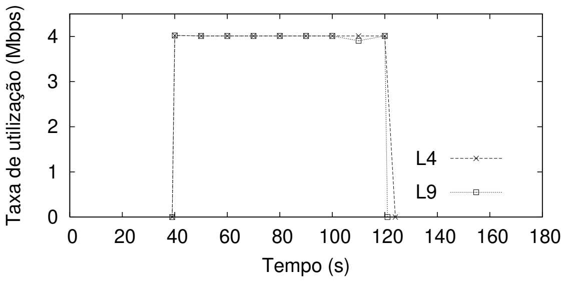
(a) Enlace L_1.



(b) Enlaces L_2 e L_7.



(c) Enlaces L_3 e L_8.

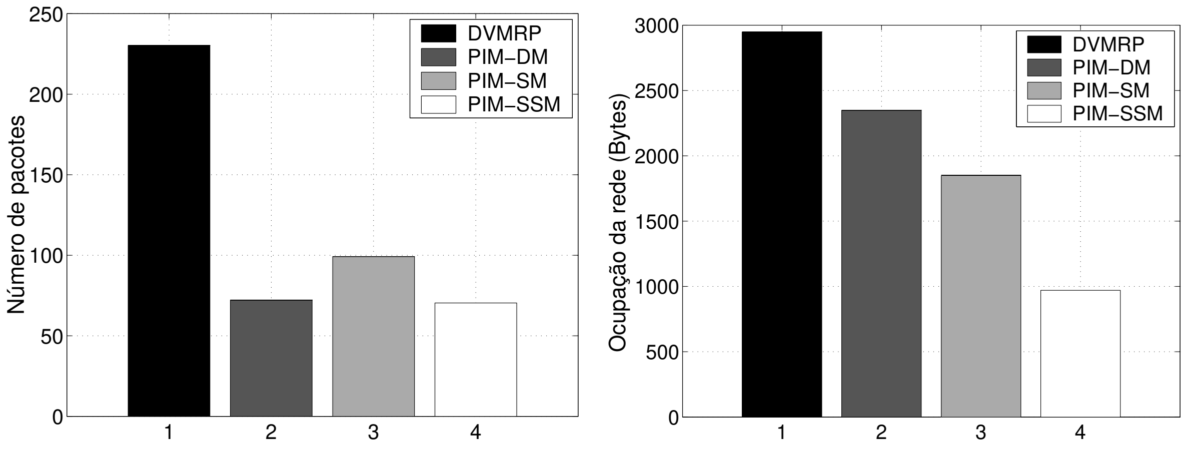


(d) Enlaces L_4 e L_9.

Figura VI.8: Ocupação dos enlaces para o PIM-SSM.

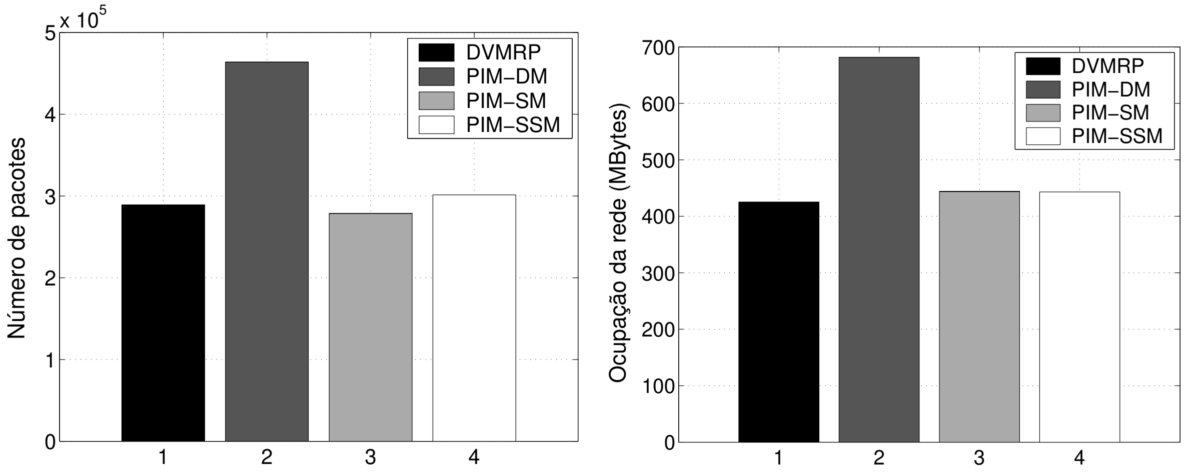
uma árvore por fonte, ao invés da árvore centrada, pois o valor da taxa de transmissão de dados da fonte era maior do que o limiar configurado nos roteadores, como pode ser visto na Tabela VI.1. O comportamento similar destes dois protocolos era esperado, já que o PIM-SSM é uma versão simplificada do PIM-SM, que só constrói árvores por fonte.

A principal diferença entre estes protocolos está no número e na carga de pacotes de controle gerados por cada um deles. As Figuras VI.9(a) e VI.9(b) mostram, respectivamente, o número total de pacotes de controle de roteamento e a carga destes pacotes em todos os enlaces da rede. Este valores correspondem à média das cinco rodadas de testes. É possível notar que o PIM-SSM envia menos mensagens de controle e que estas mensagens representam a menor carga na rede. A carga de controle gerada pelo PIM-SSM é 48,0% menor do que a gerada pelo PIM-SM. Quanto aos protocolos de inundação-e-poda, ambos geram uma maior carga de controle na rede. No entanto, por implementar o seu próprio protocolo de roteamento *unicast*, o DVMRP é o protocolo que envia o maior número de pacotes e gera a maior carga de controle na rede.



(a) Número de pacotes. (b) Ocupação da rede.

Figura VI.9: Pacotes de controle de roteamento em todos os enlaces.



(a) Número de pacotes. (b) Ocupação da rede.

Figura VI.10: Pacotes de dados em todos os enlaces.

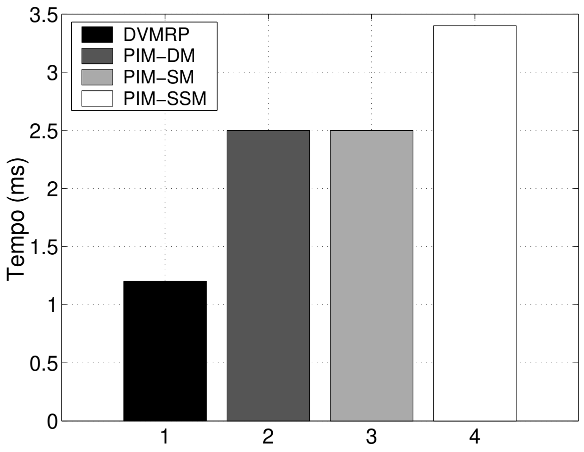
As Figuras VI.10(a) e VI.10(b) mostram o número de pacotes de dados em todos os enlaces da rede e carga gerada por estes pacotes em toda a rede. Este valores correspondem à média das cinco rodadas de testes. Por estes gráficos, pode-se comprovar o que foi observado nas Figuras VI.6(a) a VI.8(d). Nota-se que o PIM-DM, por usar um mecanismo de inundação-e-poda, apresenta uma taxa de utilização da rede 25,0% superior a do PIM-SM e a do PIM-SSM. Os dados do DVMRP não podem ser usados para fins de comparação, pois o protocolo não apresentou um funcionamento correto.

Também foi possível medir os tempos:

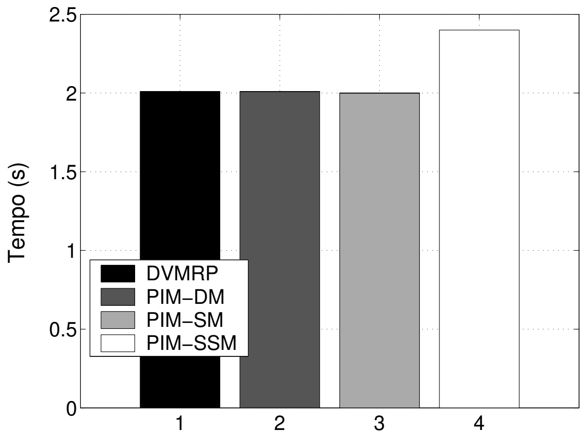
- entre o envio de uma mensagem IGMP pelo receptor, indicando o interesse no conteúdo de um grupo, e o recebimento do primeiro pacote de dados *multicast* por este receptor;
- entre o envio de uma mensagem IGMP pelo receptor, indicando o interesse em deixar um grupo, e a parada do tráfego no enlace entre o cliente e o roteador local.

As Figuras VI.11(a) e VI.11(b) mostram os valores referentes ao cliente C_2. Este cliente foi escolhido por ser o primeiro a se inscrever no grupo, ou a assinar o canal. Este valores correspondem à média das cinco rodadas de testes. Nota-se que todos os protocolos levam poucos milissegundos para começar a receber os pacotes *multicast*. Já para deixar de receber os pacotes, os clientes levam alguns segundos. O tempo entre o envio da mensagem para deixar o grupo e a parada do tráfego de dados, corresponde ao intervalo de duas mensagens de interrogação IGMP, específicas para o endereço do grupo *multicast* do qual o cliente deixou de ter interesse.

O tempo, entre o envio de uma mensagem de cancelamento da assinatura do canal e a parada do tráfego no enlace entre o cliente e o roteador local, corresponde ao tempo da troca de um canal, no serviço SSM. Este parâmetro é importante para o bom funcionamento do serviço de TV na Internet, usando o SSM. Tão logo um cliente mude de canal, haverá pacotes vindos de dois canais diferentes trafegando no enlace entre o roteador local e o cliente, o que pode provocar a perda de pacotes e a deterioração da transmissão. Este problema se torna mais grave conforme o usuário aumenta a freqüência de mudança de canais. Para evitar o congestionamento no enlace entre o roteador local e o receptor, pode-se reduzir o intervalo de tempo entre o envio de mensagens de interrogação IGMP.



(a) Tempo de entrada.



(b) Tempo de saída.

Figura VI.11: Tempos de entrada e saída de um grupo/canal *multicast*.

Porém, a carga de controle neste enlace aumentaria. Portanto, para a distribuição de TV na Internet usando o serviço SSM, será necessário o desenvolvimento de mecanismos que acelerem a troca do canal *multicast*.

Capítulo VII

Conclusão

Este trabalho teve como objetivo analisar a viabilidade dos serviços de comunicação de grupo IP Multicast e SSM (*Source-Specific Multicast*) como mecanismos de distribuição de televisão na Internet. Para isso, foi desenvolvida uma plataforma de testes, que simula um domínio típico da Internet, com suporte aos protocolos necessários para implementar os serviços IP Multicast e SSM.

A plataforma de testes é composta por um pequeno protótipo de canal de TV, por roteadores e clientes. O protótipo de emissora de TV é um único computador que tem as funções de servidor de vídeo, máquina de codificação e ilha de edição. Os roteadores são computadores pessoais rodando o sistema operacional Linux e a ferramenta de roteamento GateD. Esta ferramenta é comercial e de código fechado. No entanto, a versão usada neste trabalho é gratuita e foi obtida por meio de solicitação no sítio do fabricante na Internet.

Através da realização de experimentos práticos, foi possível comprovar a eficiência do uso de um serviço *multicast*, em termos de banda passante, para a aplicação de distribuição de TV. Os resultados mostram que à medida que o número de clientes, que requisitam o conteúdo de vídeo a uma fonte, aumenta, a taxa de utilização da rede para a distribuição *multicast* se mantém constante. Já para a transmissão ponto-a-ponto, a taxa de utilização da rede cresce linearmente com o aumento do número de clientes.

Os testes realizados também mostraram a eficiência do serviço SSM, para distribuição de aplicações de fonte única, como a TV na Internet. Por reduzir a conversação *multicast* para um-para-muitos e usar o conceito de canal como abstração de serviço, o modelo de serviço SSM elimina o problema da alocação de endereços, permite o controle de acesso ao canal e garante a integridade das aplicações. Estes são problemas característicos do serviço IP Multicast, que devido a sua complexidade, vem sendo implantado lentamente na Internet.

O desempenho de alguns dos principais protocolos de roteamento *multicast* também foi analisado. A avaliação do desempenho destes protocolos é fundamental para determinar qual o protocolo mais adequado à aplicação de distribuição de TV, uma vez que estes protocolos são os responsáveis pela criação da árvore de distribuição do tráfego *multicast*. Os protocolos analisados foram o DVMRP, o PIM-DM, o PIM-SM, usados no IP Multicast, e o PIM-SSM, usado no serviço SSM. Para avaliar o comportamento de cada protocolo, foram realizados experimentos que simularam uma transmissão de um programa de TV na Internet. As métricas usadas para comparar o desempenho dos protocolos foram a complexidade de configuração, a taxa de utilização dos enlaces da rede e o número e a carga de pacotes de controle de roteamento.

O protocolo DVMRP possui a maior carga de controle, por implementar o seu próprio protocolo de roteamento *unicast*. Além disso, o protocolo não funcionou corretamente durante os testes, uma vez que deixou de entregar o tráfego de pacotes *multicast* para um cliente que ainda era membro do grupo. Já o PIM-DM, que usa um mecanismo de inundação-e-poda para construir sua árvore de distribuição, apresenta uma taxa de utilização da rede 25,0% superior a dos protocolos de mensagens de inscrição e poda explícitas, o PIM-SM e o PIM-SSM. Estes dois protocolos possuem a menor taxa de utilização da rede. Entretanto, o PIM-SSM gera a metade da carga de controle de roteamento do PIM-SM. Além disso, o protocolo PIM-SM é o que exige o maior esforço de configuração, devido à necessidade do uso de um roteador de ponto de encontro. Por tudo isso, o PIM-SSM é o protocolo mais indicado para ser usado no serviço de distribuição de TV na Internet.

Como trabalhos futuros pretende-se continuar com desenvolvimento do protótipo de canal de TV, principalmente na produção de conteúdo e na modificação de aplicações clientes de vídeo, para que estas suportem o protocolo IGMPv3 e o serviço SSM. Também serão analisadas propostas de mecanismos para garantir o controle de acesso e a integridade para aplicações de comunicação de grupo.

Referências Bibliográficas

[1] L. H. M. K. Costa e O. C. M. B. Duarte, "Roteamento Multicast na Internet", *Mini-Curso do XXI Simpósio Brasileiro de Redes de Computadores - SBRC'2003*, pp. 228–275, maio de 2003.

[2] S. Deering, *Host Extensions for IP Multicasting.* RFC 1112, agosto de 1989.

[3] H. Holbrook e B. Cain, *Source-Specific Multicast for IP.* Work in progress, <draft-ietf-ssm-arch-04.txt>, outubro de 2003.

[4] H. W. Holbrook e D. R. Cheriton, "IP multicast channels: EXPRESS support for large-scale single-source applications", in *ACM SIGCOMM'99*, setembro de 1999.

[5] A. S. Tanenbaum, *Computer Networks.* Prentice Hall Inc, terceira ed., 1996.

[6] I. M. Moraes, M. D. D. Bicudo, K. V. Cardoso, S. V. Vasconcellos, J. F. de Rezende e O. C. M. B. Duarte, "Desenvolvimento de um Ambiente de Testes com Suporte à Qualidade de Serviço para a Transmissão de Vídeo Digital", in *WRNP2'2003 - IV Workshop WRNP2*, maio de 2003.

[7] P. A. da Silva Gonçalves, J. F. de Rezende e O. C. M. B. Duarte, "Um serviço ativo de distribuição de vídeo multiponto", in *XVIII Simpósio Brasileiro de Redes de Computadores SBRC'2000*, pp. 211–226, maio de 2000.

[8] S. Y. Cheung, M. H. Ammar e X. Li, "On the use of destination set grouping to improve fairness in multicast video distribution", in *IEEE INFOCOM'96*, pp. 553–560, 1996.

[9] E. Amir, S. McCanne e H. Zhang, "An application level video gateway", *In Proc. ACM Multimedia '95, San Francisco, CA*, 1995.

[10] S. McCanne, V. Jacobson e M. Vetterli, "Receiver-driven layered multicast", in *ACM SIGCOMM'96*, outubro de 1996.

[11] B. Vickers, C. Albuquerque e T. Suda, "Adaptive multicast of multi-layered video: Rate-Based and Credit-Based Approaches", in *IEEE INFOCOM'98*, fevereiro de 1998.

[12] H. Schulzrinne, S. Casner, R. Frederick e V. Jacobson, *RTP: A Transport Protocol for Real-Time Applications.* RFC 1889, janeiro de 1996.

[13] H. Schulzrinne, A. Rao e R. Lanphier, *Real Time Streaming Protocol (RTSP).* RFC 2326, abril de 1998.

[14] S. D. dos Reis, A. Proença e D. P. Júnior, "Modelo de Négocio: um exercício conceitual sobre o caso TV aberta x TV por assinatura", in *XXIII Encontro Nacional de Engenharia de Produção – ENEGEP'2003*, outubro de 2003.

[15] D. Waterman, "Internet TV: Business Models and Program Content", in *Research Conference on Information, Communication and Internet Policy*, setembro de 2001.

[16] W. Fenner, *Internet Group Management Protocol, Version 2*, novembro de 1997. RFC 2236.

[17] C. Huitema, *Routing in the Internet.* Prentice Hall Inc, segunda ed., 2000.

[18] B. Cain, S. Deering, I. Kouvelas, B. Fenner e A. Thyagarajan, *Internet Group Management Protocol, Version 3*, outubro de 2000. RFC 3376.

[19] A. Adams, J. Nicholas e W. Siadak, *Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised).* Work in progress, <draft-ietf-pim-dm-new-v2-04.txt>, setembro de 2003.

[20] B. Fenner, M. Handley, H. Holbrook e I. Kouvelas, *Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)*. Work in progress, <draft-ietf-pim-sm-v2-new-08.txt>, outubro de 2003.

[21] S. Bhattacharyya, *An Overview of Source-Specific Multicast (SSM)*. RFC 3569, julho de 2003.

[22] H. Sandick e B. Cain, *PIM-SM rules for Support of Single-Source Multicast*. Work in progress, <draft-sandick-pimsm-ssmrules-00.txt>, março de 2000.

[23] M. Handley, C. Perkins e E. Whelan, *Session Announcement Protocol*. RFC 3569, outubro de 2000.

[24] "Darwin Streaming Server", janeiro de 2004. `http://developer.apple.com/darwin/projects/streaming/`.

[25] "MPEG4IP", dezembro de 2003. `http://mpeg4ip.sourceforge.net`.

[26] "VideoLAN", fevereiro de 2004. `http://www.videolan.org`.

[27] "Live.com Streaming Media", fevereiro de 2004. `http://www.live.com/liveMedia`.

[28] "Openmash project", fevereiro de 2004. `http://www.openmash.org`.

[29] V. Kulkarni, "IGMPv3 Linux 2.4.20 Patch", fevereiro de 2004. `http://people.internet2.edu/~admytren/convena/files/igmpv3_2_4_20.patch.gz`.

[30] "NextHop Technologies. GateD 10.0 Command Refernce", outubro de 2003. `http://www.nexthop.com/products/datasheet/cli-config.pdf`.

[31] "Iperf - The TCP/UDP Bandwidth Measurement Tool", fevereiro de 2004. `http://dast.nlanr.net/Projects/Iperf/`.

[32] L. Lhotka, "SSM-enabled applications", fevereiro de 2004. `http://www.cesnet.cz/tf-ngn/multicast/ssm/`.

[33] D. Mackie, "Video streaming and MPEG4IP", in *Cisco Technology Center*, 2001.

Apêndice A

Arquivos

Este apêndice contém exemplos dos arquivos de configuração da ferramenta GateD, usada nos roteadores da plataforma de testes, e também dos *scripts*, usados para simular uma transmissão de TV na Internet.

A.1 Configuração dos Roteadores

A ferramenta GateD armazena todas as suas configurações no arquivo `gated.conf`. A seguir, são apresentados alguns exemplos de arquivos de configuração para os diferentes roteadores e protocolos usados na plataforma de testes.

A.1.1 Protocolo DVMRP

O arquivo mostrado a seguir corresponde à configuração do roteador R_1, quando este executa o protocolo DVMRP.

```
interfaces {
  interface
    "eth0"
    disable
  ;
```

```
};
dvmrp on {
  interface
    "eth2"
  ;
  interface
    "eth3"
  ;
};
igmpv3 on {
  interface
    "eth1"
  ;
};
ospf on {
  instance
    1
  { area
      0.0.0.10
    {
        match 192.168.10
        wildcard 0.0.0.255
      ;
      match 192.168.11
      wildcard 0.0.0.255
      ;
      match 192.168.21
      wildcard 0.0.0.255
      ;
    }
    ; }
  ;
  floating-interface
    "eth1"
  { passive on
    ;
  }
  ;
  floating-interface
    "eth2"
  ;
  floating-interface
    "eth3"
  ;
};
```

A.1.2 Protocolo PIM-DM

O arquivo mostrado a seguir corresponde à configuração do roteador R_2, quando este executa o protocolo PIM-DM.

```
interfaces {
  interface
    "eth0.2"
    disable
  ;
};
igmpv3 on {
  interface
    "eth0.14"
  ;
};
pim on {
  interface
    "eth0"
  ;
  interface
    "eth1"
  { mode dense
    ; }
  ;
  interface
    "eth2"
  { mode dense
    ; }
  ;
  interface
    "eth3"
  { mode dense
    ; }
  ;
};
ospf on {
  instance
    1
  { defaults
    { multicast-rib on
      ; }
    ;
    area
```

```
      0.0.0.10
    {
        match 192.168.11
        wildcard 0.0.0.255
      ;
      match 192.168.12
      wildcard 0.0.0.255
      ;
      match 192.168.14
      wildcard 0.0.0.255
      ;
      match 192.168.30
      wildcard 0.0.0.255
      ;
    }
    ; }
  ;
  floating-interface
    "eth0.14"
  { passive on
    ;
  }
  ;
  floating-interface
    "eth1"
  ;
  floating-interface
    "eth2"
  ;
  floating-interface
    "eth3"
  ;
};
```

A.1.3 Protocolo PIM-SM

O arquivo mostrado a seguir corresponde à configuração do roteador R_3, quando este executa o protocolo PIM-SM.

```
interfaces {
  interface
    "eth0.2"
```

```
    disable
  ;
};
igmpv3 on {
  interface
    "eth0.13"
  ;
};
pim on {
  interface
    "eth1"
  { mode sparse
    ; }
  ;
  interface
    "eth2"
  { mode sparse
    ; }
  ;
  threshold 4000000000
  ;
  threshold-dr 4000000000
  ;
  threshold-rp 4000000000
  ;
  static-rp
    224
    mask-length 4
    192.169.11.37
  ;
};
ospf on {
  instance
    1
  { defaults
    { multicast-rib on
      ; }
    ;
    area
      0.0.0.10
    {
        match 192.168.12
        wildcard 0.0.0.255
      ;
      match 192.168.13
```

```
      wildcard 0.0.0.255
      ;
      match 192.168.22
      wildcard 0.0.0.255
      ;
    }
    ;
  ;
  floating-interface
    "eth0.13"
  { passive on
    ;
  }
  ;
  floating-interface
    "eth1"
  ;
  floating-interface
    "eth2"
  ;
};
```

A.1.4 Protocolo PIM-SSM

O arquivo mostrado a seguir corresponde à configuração do roteador R_4, quando este executa o protocolo PIM-SSM.

```
igmpv3 on {
  interface
    "eth0"
  ;
};
pim on {
  interface
    "eth0"
  ;
  interface
    "eth1"
  { mode ssm
    ; }
  ;
  interface
```

```
    "eth2"
  { mode ssm
    ; }
  ;
  interface
    "eth3"
  { mode ssm
    ; }
  ;
};
ospf on {
  instance
    1
  { defaults
    { multicast-rib on
      ; }
    ;
    area
      0.0.0.10
    {
        match 192.168.15
        wildcard 0.0.0.255
      ;
      match 192.168.21
      wildcard 0.0.0.255
      ;
      match 192.168.22
      wildcard 0.0.0.255
      ;
      match 192.168.30
      wildcard 0.0.0.255
      ;
    }
    ;
  ;
  floating-interface
    "eth0"
  { passive on
    ;
  }
  ;
  floating-interface
    "eth1"
  ;
  floating-interface
```

```
    "eth2"
  ;
  floating-interface
    "eth3"
  {
  ;
};
```

A.2 *Scripts*

Os *scripts* são pequenos programas escritos em linguagem de *shell*. Neste trabalho, eles foram usados para simular o cenário dos testes.

A.2.1 Simulação

Para simular a transmissão de um programa de TV, foram criados *scripts* correspondentes ao comportamento do servidor de vídeo e dos clientes de vídeo. A seguir são apresentados os *scripts* para o servidor de vídeo e para o cliente C_2.

Servidor de Vídeo

```
#!/bin/bash
CLIENT1="icarai"
CLIENT2="lumiar"
CLIENT3="grumari"
FILEDIR="/root"
FLAGDIR1="/net/$CLIENT1/home"
FLAGDIR2="/net/$CLIENT2/home"
FLAGDIR3="/net/$CLIENT3/home"
HOMEDIR="/net/grumari/home/igor"
IPERF_FILE=""$FILEDIR/iperf-videoserver-$4-$5
TIME_IN="10"
USE="Uso:$0 <grupo> <tempo (s)> <banda (M)> <protocolo> <rodada>"

#Testando os parâmetros
[ $# -ne 5 ] && echo $USE && exit 1
```

```
#Sincronizando o relógio
/usr/sbin/ntpdate 146.164.69.28

#Removendo os flags
#rm $FLAGDIR2/flag
#rm $FLAGDIR3/flag
#rm $FLAGDIR4/flag

#Executando o iperf
iperf -u -c $1 -T 16 -t $2 -i 1 -b $3 -p 10000 >> $IPERF_FILE &
id_iperf=$!
[ $? -ne 0 ] && echo "Error: iperf" && exit 1
echo "Running iperf..." && date

#Cliente 1
sleep $TIME_IN
date
touch $FLAGDIR1/flag
[ $? -ne 0 ] && echo "Error: touch - $CLIENT1" && exit 1

#Cliente 2
sleep $TIME_IN
date
touch $FLAGDIR2/flag
[ $? -ne 0 ] && echo "Error: touch - $CLIENT2" && exit 1

#Cliente 3
sleep $TIME_IN
date
touch $FLAGDIR3/flag
[ $? -ne 0 ] && echo "Error: touch - $CLIENT3" && exit 1

# Teste se existe o arquivo para movê-lo
mv -f $FILEDIR/iperf-videoserver-$4-$5 $HOMEDIR/$4/$5

date
echo "fim"
```

Clientes

```
#!/bin/bash
HOMEDIR="/net/grumari/home/igor"
FLAGDIR="/home"
```

```
FILEDIR="/root"
CLIENT="c14"
TIME_IN="10"
INTERVAL="0.5"
USE="Uso:$0 <grupo> <protocolo> <rodada> <tempo (s)>"

#Testando os parâmetros
[ $# -ne 4 ] && echo $USE && exit 1

#Sincronizando o relógio
/usr/sbin/ntpdate 146.164.69.28

#Verifica a existência do flag de INTERVAL em INTERVAL segundos
while [ ! -r $FLAGDIR/flag ];
do
 sleep $INTERVAL
done

#Quando o flag é criado...
if [ $2 = ssm ]; then
   echo 1
     ssmrec-static -p 10000 -s 192.168.10.3 $1 >> $CLIENT_FILE &
else
  echo 2
     iperf -s -u -B $1 -i 1 >> $CLIENT_FILE &
fi

# Tempo em que o cliente recebe o "vídeo"
sleep $4

#Cliente deixa o grupo
if [ $2 = ssm ]; then
  killall ssmrec-static
else
  killall iperf
fi

# Movendo o arquivo
mv -f $FILEDIR/$CLIENT-$2-$3 $HOMEDIR/$2/$3

date
rm $FLAGDIR/flag
echo fim
```

A.2.2 Medição

Durante a simulação do programa de TV, é necessário monitorar os enlaces da plataforma de testes. A seguir é apresentado um exemplo de *script* usado para monitorar as interfaces dos roteadores, neste caso R_1.

```
#!/bin/bash

HOMEDIR="/net/grumari/home/igor"
FILEDIR="/root"
TEMPFILE="./tmp"
USE="Uso:$0 <protocolo> <rodada> <tempo (s)>"

# Testando os parâmetros
[ $# -ne 3 ] && echo $USE && exit 1

# Sincronizando o relógio
/usr/sbin/ntpdate 146.164.69.28

# Monitorando as interfaces durante o tempo passado como parâmetro
tcpdump -vvv -i eth1 -U root >> $TEMFILE-eth1 &
tcpdump -vvv -i eth3 -U root >> $TEMPFILE-eth3 &
sleep $3
killall tcpdump

# Criando os arquivos de acordo com o protocolo
# Pacotes UDP vindos da fonte
awk '($2 == /videoserver/){ print }' ./tmp-eth1 >> $FILEDIR/$1-l10-udp-$2
awk '($2 == /videoserver/){ print }' ./tmp-eth3 >> $FILEDIR/$1-l21-udp-$2

# Protocolos de roteamento multicast
if [ $1 -eq dvmrp ]; then
  awk '($5 == /dvmrp/){ print }' ./tmp-eth3 >> $FILEDIR/$1-l21-ctrl-$2
else
  awk '($5 == /pim/){ print }' ./tmp-eth3 >> $FILEDIR/$1-l21-ctrl-$2
fi

# Protocolo de roteamento unicast
awk '($5 == /OSPF/){ print }' ./tmp-eth3 >> $FILEDIR/$1-l21-ospf-$2

# IGMP
awk '($5 == /igmp/){ print }' ./tmp-eth1 >> $FILEDIR/$1-l10-igmp-$2
```

```
# Movendo os arquivos
mv -f $FILEDIR/$1-l10-udp-$2 $HOMEDIR/$1/$2
mv -f $FILEDIR/$1-l21-udp-$2 $HOMEDIR/$1/$2
mv -f $FILEDIR/$1-l21-ctrl-$2 $HOMEDIR/$1/$2
mv -f $FILEDIR/$1-l21-ospf-$2 $HOMEDIR/$1/$2
mv -f $FILEDIR/$1-l10-igmp-$2 $HOMEDIR/$1/$2

date
rm ./TEMPFILE-eth1
rm ./TEMPFILE-eth3
echo fim
```