Escalabilidade no Roteamento Baseado em Requisitos de Qualidade de Serviço

Luís Henrique Maciel Kosmalski Costa
luish@gta.ufrj.br

Otto Carlos Muniz Bandeira Duarte
otto@gta.ufrj.br

Grupo de Teleinformática e Automação – GTA
COPPE/EE - Programa de Engenharia Elétrica/Departamento de Eletrônica
Universidade Federal do Rio de Janeiro
Caixa Postal 68504 - CEP 21945-970 - Rio de Janeiro - RJ - Brasil

Resumo

O roteamento baseado em requisitos de Qualidade de Serviço (QoS) é um componente fundamental na definição de uma nova arquitetura de serviços integrados para a Internet. Vários aspectos merecem atenção, como os requisitos deste tipo de roteamento, que tipo de arquitetura pode ser utilizada na prática, quais as considerações técnicas e políticas que devem ser feitas numa arquitetura de roteamento Internet baseada em QoS. Uma das principais características da implementação do roteamento baseado em QoS é a garantia da escalabilidade deste esquema. Neste trabalho, é feita a proposta de um algoritmo aproximativo para o cálculo de rotas com três restrições: atraso, faixa de transmissão e probabilidade de perda.

Abstract

Quality of Service (QoS) based routing is a fundamental part of the definition of a new integrated services architecture for the Internet. Several questions deserve attention, as the requirements for this routing model, what kind of architecture is practical, what are the technical and political issues that arise in implementing a QoS-based Internet routing architecture. One of the main characteristics on the implementation of QoS-based routing is guaranteeing scalability for that scheme. In this work an approximate algorithm is presented for the computation of paths with three requirements: delay, transmission bandwidth and loss probability.

1 Introdução

A Internet vem sendo usada cada vez mais no transporte de informação multimídia, como imagens, voz e vídeo. Com o aumento da capacidade de processamento dos computadores pessoais e o surgimento de novas aplicações multimídias, a Internet tende a receber uma maior quantidade deste tipo de tráfego. Este tipo de utilização, no entanto, não é o uso para o qual o protocolo IP (*Internet Protocol*) foi projetado, há mais de vinte anos. A arquitetura de rede da Internet foi projetada para o envio da informação usando a disciplina de serviço de melhor esforço, ou seja, sem quaisquer garantias de vazão ou de atraso.

Para o suporte de tráfego de tempo real na Internet, algumas novas características e serviços devem ser adicionados. Novas classes de serviço precisam ser oferecidas, além do envio de melhor esforço, para prover alguma garantia de atraso e vazão às aplicações. É necessário um mecanismo que permita às aplicações sinalizar seus requisitos de QoS à rede, e mecanismos de gerenciamento de tráfego nas estações e nos roteadores.

O IETF (*Internet Engineering Task Force*) vem trabalhando nas duas primeiras questões, a definição de novas classes de serviço [1] e um protocolo de reserva de recursos, o RSVP (*Resource ReSerVation Protocol*) [2, 3, 4], que as aplicações poderão utilizar para solicitar à rede a QoS desejada. Outra área que merece atenção é a definição de uma arquitetura de roteamento que considere os requisitos de QoS das aplicações. Foi criado um grupo de trabalho específico dentro do IETF para tratar das questões relacionadas ao roteamento baseado em QoS, através das principais características que uma tal arquitetura deve possuir.

Um dos pontos principais de uma arquitetura de roteamento baseado em requisitos de QoS é a garantia de sua escalabilidade. Ou seja, o esquema adotado deve possuir uma complexidade tal que não se inviabilize com o crescimento da quantidade de nós da rede cresça. Esta característica toma importância cada vez maior dado o ritmo de crescimento atual da Internet. Dentro da garantia de escalabilidade do protocolo de roteamento, dois pontos se destacam: a complexidade do algoritmo implementado para a descoberta de rotas e o tamanho das tabelas de roteamento utilizadas. A definição de algoritmos para o cálculo de rotas com múltiplas métricas vem recebendo grande atenção.

Na seção 2 deste trabalho, são analisados os principais requisitos que o roteamento baseado em QoS deve respeitar. Na seção 3 são apresentadas as principais características relacionadas com os algoritmos de roteamento, considerando requisitos de QoS comuns em grande parte das aplicações. A seção 4 apresenta uma proposta para a construção de uma arquitetura de roteamento escalável, através da definição de algoritmos para o cálculo de rotas com três restrições. Na seção 5 são feitas algumas considerações finais.

2 Requisitos do Protocolo de Roteamento

Considerando a organização da Internet, a natureza do tráfego gerado, as classes de serviço que devem ser providas e os requisitos dos usuários, podem ser definidos diversos pré-requisitos ao roteamento baseado em reservas de recursos. Para uma melhor organização, a seguir são apresentados diversos requisitos, divididos nos grupos principais: métricas e rotas, robustez, eficiência, prioridades, inter-domínios, políticas e escalabilidade. Outros requisitos, como o suporte a mobilidade, podem ser definidos [5].

Quanto a métricas e rotas, é desejável que o protocolo de roteamento suporte múltiplas métricas para os caminhos (atraso, faixa de transmissão, variação de atraso etc.). A definição das métricas necessárias deve ser realizada baseada nas classes de serviço suportadas [6, 7]. O protocolo deve ser capaz de calcular dinamicamente rotas satisfazendo múltiplas restrições. Deve oferecer suporte de QoS para vários tipos de fluxos, incluindo ponto-a-ponto (*unicast*) e ponto-a-multiponto (*multicast*) e, além disso, deve suportar receptores heterogêneos em termos dos requisitos de QoS.

O protocolo deve ser robusto, ou seja, possuir as habilidades de detectar e responder a mudanças nas capacidades de QoS dos enlaces e de continuar o roteamento mesmo diante de falhas múltiplas na rede. Em resumo, o protocolo deve ser adaptável às variações da rede.

Em se tratando de eficiência, deseja-se que o protocolo seja capaz de rotear fluxos de curta duração com a menor sobrecarga de processamento possível, além de utilizar os recursos da rede de forma eficiente, através de balanceamento da carga, e controles de admissão de chamadas globais. A sobrecarga de processamento devida ao roteamento deve ser mínima. Deve ser possível o controle dos recursos, ou seja, a capacidade de limitar o consumo de recursos da rede pelas diferentes classes de tráfego.

Deve existir um mecanismo para estabelecer prioridades de transmissão entre os fluxos e para permitir que tráfegos de alta prioridade sejam estabelecidos com precedência em relação aos menos prioritários.

Os requisitos para o roteamento entre domínios diferentes constituem um aspecto importante que tem gerado bastante discussão [8]. Mais especificamente, a natureza da informação trocada na fronteira entre dois domínios administrativos deve ser determinada. Se a troca de informação for pequena, o roteamento inter-domínio pode ser relativamente simples. Em todo caso, a capacidade de introdução de restrições de acordo com políticas no roteamento do fluxo de informação na vizinhança entre duas organizações é indispensável.

O protocolo deve ser escalável, pré-requisito indispensável quando se considera a dimensão de uma rede como a Internet. O algoritmo utilizado deve ser capaz de lidar com o aumento do número de nós e de enlaces da rede sem se tornar inviável em termos da capacidade de processamento das máquinas, assim como de sua capacidade de memória. Ou seja, as tabelas de roteamento mantidas não podem crescer descontroladamente com o tamanho das redes. Nas seções seguintes, alguns dos aspectos relacionados com o protocolo de roteamento são explorados.

2.1 Determinação da qualidade de serviço e reserva de recursos

O Grupo de Trabalho em Serviços Integrados [1, 9, 10] do IETF vem se concentrando no gerenciamento local de recursos nos roteadores. A noção de controle de admissão [11] foi introduzida, permitindo ao roteador a decisão de se um novo fluxo deve ser aceito ou rejeitado, baseado na disponibilidade local de recursos. O controle de admissão requer o conhecimento da QoS disponível em todos os enlaces conectados ao roteador. É uma questão em aberto, no entanto, como os valores de QoS são calculados para enlaces de diferentes tipos. Esta questão está no domínio do Grupo de Trabalho ISSLL (*Integrated Services over Specific Link Layers*) do IETF.

Este problema ocorre, por exemplo, no caso de um roteador conectado a uma rede não *broadcast* de múltiplo acesso, como a rede ATM (*Asynchronous Transfer Mode*). Nesta situação, o roteador deverá ter diversas escolhas para o próximo nó, com caminhos correspondentes através da rede ATM, cada um dos quais possivelmente com diferente disponibilidade de QoS. As questões que se impõem são como o roteador pode determinar quais são as opções de roteamento através da rede ATM, qual é a disponibilidade de recursos em cada um destes caminhos, e que valores de QoS devem ser informados pelo roteador à rede ATM, no caso do roteamento baseado em reserva de recursos ser implementado na Internet. Como uma ilustração destes problemas, os padrões atualmente definidos para o roteamento IP sobre ATM, como o NHRP (*Next Hop Routing Protocol*), permitem a seleção de um único ponto de saída da rede ATM para um datagrama IP. Com requisitos de QoS nos fluxos IP, pode não existir um caminho entre o roteador de entrada e o roteador de saída definido pelo NHRP que suporte a exigência de QoS.

Um problema relacionado é a reserva de recursos numa rede *broadcast*, como a rede Ethernet. Como o acesso à rede é distribuído, uma coordenação entre roteadores e estações torna-se indispensável para uma reserva de recursos.

2.2 Propagação da informação

A arquitetura de roteamento baseada no estado do enlace requer que os roteadores informem através de difusão o *status* dos recursos locais (como a faixa de transmissão disponível e os atrasos medidos) e as informações relacionadas com a topologia local para todos os outros roteadores da rede. A difusão da informação de *status* de um roteador para os outros tende a ser um processo dispendioso em termos de sobrecarga do sistema de comunicação e das capacidades de processamento. Até então, o roteamento intra-domínio tem sido baseado em métricas fixas para os enlaces. Neste contexto, a eficiência na propagação da informação não era um pré-requisito. Por esse motivo, os algoritmos de roteamento como o implementado pelo OSPF (*Open Shortest Path First*) utilizam-se de inundação com um período de *refresh* como forma de propagar as atualizações.

Para o roteamento baseado em QoS eficiente, no entanto, mecanismos de inundação não são adequados para a distribuição da informação. Técnicas alternativas para reduzir a sobrecarga da difusão, como o envio baseado em árvores, foram propostas. Além disso, a informação do *status* dos enlaces é em muitos casos volátil, ou seja, seus valores podem mudar significativamente em curtos intervalos de tempo com a admissão de fluxos de tempo real. A maneira como essa informação é quantificada de forma a diminuir a freqüência das atualizações é uma questão importante. Quanto menos precisa for a informação, menor a chance de um roteador fonte encontrar um caminho ideal. Por outro lado, os custos da manutenção de informação muito acurada em cada roteador podem ser muito altos.

2.3 Prioridades entre os fluxos

No caso de existirem fluxos críticos que precisem ter uma maior prioridade que outros fluxos, um mecanismo deve ser implementado na rede para reconhecer os níveis de prioridade. Assume-se que o RSVP possa ser utilizado para sinalizar a prioridade de um fluxo. Existem dois aspectos na priorização entre fluxos. O primeiro, diz respeito à maneira como diferentes usuários poderão controlar a prioridade de seus fluxos. A rede deve ser capaz de conferir se um determinado fluxo tem o direito de utilizar a prioridade requisitada. Em segundo lugar, o mecanismo de roteamento deve assegurar que o caminho com uma determinada QoS será encontrado com uma probabilidade proporcional ao nível de prioridade do fluxo em questão. Ou seja, para uma determinada carga da rede, um fluxo com maior prioridade deve ter mais chances de ser aceito que um fluxo com os mesmos requisitos de QoS, mas com uma menor prioridade.

O mecanismo de roteamento para implementação de prioridades entre os fluxos deve ser baseado em preempção combinada com o re-roteamento dinâmico dos fluxos de baixa prioridade. Num primeiro momento, o roteador deve procurar caminhos com a disponibilidade de recursos necessária para o novo fluxo ser aceito. Não havendo tal caminho, o mecanismo deve tentar um caminho onde recursos adequados possam ser alocados pelo remanejamento de fluxos de menor prioridade. Em último caso, seria usada a preempção dos recursos alocados para os fluxos de menor prioridade. Os roteadores devem tentar o re-roteamento dos fluxos atingidos pela preempção utilizando o mesmo procedimento. O esquema de roteamento considerando prioridades pode, portanto, se tornar complexo.

2.4 Controle de recursos

Dada a existência de múltiplas classes de serviço, é necessário planejar uma rede para suportar as demandas de tráfego previstas para cada classe. Neste sentido, um administrador de rede pode particionar logicamente os recursos dos roteadores e dos enlaces entre as diversas classes de serviço. Este particionamento rigoroso, porém, resultará em uso ineficiente dos recursos da rede, pois durante determinados períodos de tempo pode haver carga excessiva de tráfego de uma classe de serviço e ociosidade de recursos em outra classe. É desejável, portanto, permitir que recursos não utilizados sejam alocados a tráfegos que necessitem. Este compartilhamento deve ser feito de forma controlada, de maneira a impedir que tráfego de uma determinada classe de serviço possa ter acesso a mais recursos do que fora projetada a rede para esta classe por períodos de tempo prolongados.

Esta é uma tarefa do mecanismo de roteamento. Pode ser realizada através de preempção, de maneira não contraditória à atribuição de prioridades dos fluxos de tráfego ativos. O projeto de um esquema de compartilhamento de recursos apropriado e sua incorporação ao roteamento baseado em QoS são importantes. Em particular, as sobrecargas impostas pela implementação do particionamento lógico de recursos e compartilhamento dinâmico é uma questão a ser estudada.

2.5 Fluxos de curta duração

O roteamento baseado em requisitos de QoS incorpora determinadas sobrecargas de processamento durante o estabelecimento do fluxo, por exemplo, no cálculo de uma rota a partir de uma fonte. A questão que se coloca é se este tempo gasto a mais é desproporcional ao tempo de duração das sessões. Em geral, o problema se torna mais grave em redes com circuitos virtuais, como a rede ATM, onde vários circuitos virtuais chaveados (*Switched Virtual Circuit* - SVC) de curta duração resultam em maiores sobrecargas de estabelecimento de conexão. Mesmo sem o roteamento baseado em QoS, as sobrecargas devidas ao estabelecimento de fluxo através do RSVP são adicionadas em cada sessão. A diminuição da sobrecarga devida ao roteamento no estabelecimento de fluxos deve ser buscada, uma das formas possíveis pode ser a utilização de *caches* com as rotas recentemente computadas, de maneira que novos fluxos para os mesmos destinos não exijam novo cálculo dos caminhos.

2.6 Algoritmos para aceitação de fluxos ponto-a-ponto

O cálculo de um caminho possível para um fluxo pode ser feito de maneira simples: o roteador elimina primeiramente todos os nós e enlaces impraticáveis, ou seja, os que não podem dar suporte aos requisitos de QoS determinados, da sua representação local da topologia da rede e encontra um caminho fim-a-fim na rede remanescente. Porém este tipo de algoritmo resulta na aceitação de um número menor de fluxos do que seria possível de outra forma. Este tipo de admissão de fluxos corresponde ao trabalho realizado pelo Grupo de Trabalho em Serviços Integrados do IETF; neste caso, um fluxo passa pelo controle de admissão à medida em que é aceito por cada roteador ao longo do caminho. Enquanto este tipo de controle é necessário para administrar os recursos em um roteador, não leva em conta qualquer visão global do consumo dos recursos pelos diversos fluxos individuais. A eficiência no roteamento deve ser obtida através de uma camada de controle de admissão a mais. Este controle de admissão de mais alto nível pode considerar o consumo de recursos de cada fluxo em relação aos recursos disponíveis ao longo de um caminho de forma a determinar se a aceitação deste fluxo é desejável globalmente. Sendo assim, um fluxo individual pode ser rejeitado mesmo se existir um caminho possível para roteá-lo, se a aceitação deste fluxo puder resultar num declínio geral do número de fluxos carregados pela rede. A determinação desta camada de controle de admissão de mais alto nível, com justiça para todos os fluxos que desejem utilizar-se da rede, é uma área de interesse atual.

O cálculo de rotas sujeito a múltiplas restrições de QoS em um fluxo é um problema complexo. Para algumas combinações de requisitos de QoS, o problema é NP-completo [12]. No entanto, alguns algoritmos foram propostos considerando requisitos de faixa de transmissão e de atraso, podendo ser incorporados aos mecanismos de roteamento. Na seção seguinte são apresentados três destes algoritmos.

2.7 Roteamento baseado em QoS para fluxos ponto-a-multiponto

O cálculo de rotas para a acomodação de fluxos ponto-a-multiponto com requisitos de QoS é um problema delicado, especialmente com a noção de um nível de controle de admissão mais alto adicionado. O dinamismo permitido pelo IP *Multicast* no grupo de receptores, aliado à heterogeneidade dos receptores complicam a questão. Considerando um esquema do tipo adotado pelo MOSPF (*Multicast* OSPF) [13, 14] a principal dificuldade diz respeito à escalabilidade. Para considerar QoS, o cálculo do caminho *multicast* no roteador deve levar em conta conhecimento não apenas da identidade das sub-redes onde membros do grupo estão presentes, mas também a identificação dos ramos na árvore existente. Ou seja, nesta versão do MOSPF, os roteadores precisam manter informação específica das fontes para cada grupo. Mudanças no grupo de destinos ou reservas realizadas por receptores resultam em informação sobre múltiplas árvores sendo distribuídas para todos os roteadores na rede.

Uma vez que a informação necessária está disponível, existem diversos algoritmos baseados na heurística para modificar incrementalmente árvores de distribuição existentes [12]. O cálculo de árvores compartilhadas baseadas no estilo de reserva compartilhado [2, 3, 4], no entanto, requer novos algoritmos.

3 Algoritmos de Roteamento

Os esquemas de roteamento pela origem (*source routing*) e roteamento salto-a-salto (*hop-by-hop routing*) estão presentes em grande parte das arquiteturas de roteamento para redes de transmissão de dados. O roteamento salto-a-salto é a forma mais comum de roteamento de propósito geral nas redes atuais, enquanto o roteamento pela origem é utilizado principalmente para o diagnóstico da rede e para realização de rotas com políticas especiais.

No esquema de roteamento pela origem, a rota de envio é calculada sob demanda na fonte da transmissão e listada explicitamente no cabeçalho do pacote. Os pacotes são encaminhados de acordo com a informação em seu cabeçalho. Uma vez que o cálculo é realizado para cada requisição, de maneira centralizada, o roteamento pela origem é flexível. No entanto, a fonte necessita de informação relativa a todos os enlaces envolvidos no caminho, e os pacotes possuem um cabeçalho longo. Existe também um tempo de cálculo durante a inicialização.

Já no roteamento salto-a-salto, os pacotes são encaminhados salto a salto a cada roteador. Cada nó possui uma tabela de roteamento com o próximo salto para todos os destinos, sendo em geral esta tabela atualizada periodicamente em resposta a mudanças na topologia da rede. Quando um pacote chega a um roteador, este precisa apenas checar na tabela qual a porta de saída para o destino desejado. Em relação ao roteamento pela origem, o cabeçalho dos pacotes pode ser menor, uma vez que não carrega a descrição completa da rota a ser seguida. O roteamento salto-a-salto pode ser implementado através de algoritmos completamente distribuídos.

Estes dois esquemas de roteamento possuem características que se encaixam nas necessidades do roteamento baseado em requisitos de QoS. Os requisitos de QoS das aplicações são bastante diversos, sendo difícil especificar um conjunto de parâmetros que sirva de base para a maioria das aplicações. Deste ponto de vista, o roteamento pela origem seria bastante apropriado, uma vez que o cálculo de rotas é realizado de acordo com cada fluxo, sob demanda. Por outro lado, o roteamento salto-a-salto permite a implementação de algoritmos distribuídos e possui as vantagens de menores sobrecarga de processamento e tempo gasto na inicialização.

Nesta seção, primeiramente é feita uma análise dos fatores de decisão na escolha da métrica ou conjunto de métricas adequado para a implementação de protocolos de roteamento baseados em QoS. Em seguida, são discutidas diferentes maneiras de combinarem-se as métricas. Esta combinação pode ser feita através da criação de uma medida que englobe as diferentes métricas, ou mediante a análise das métricas separadamente, por meio de algoritmos específicos.

3.1 Análise das métricas possíveis

A métrica escolhida pelo roteamento define basicamente a forma como a rede será representada. Esta decisão influirá não somente na complexidade do algoritmo de roteamento a ser implementado, mas também no conjunto de requisitos de QoS que tal esquema será capaz de suportar. Diversos aspectos devem ser levados em conta na escolha da métrica ou do conjunto de métricas que constituirão o fator de decisão de roteamento [12, 15].

Qualquer métrica eleita deve permitir algoritmos eficientes para o cálculo das rotas, de forma a garantir a escalabilidade do protocolo de roteamento, indispensável quando leva-se em consideração o tamanho de uma rede como a Internet. De preferência, a complexidade dos algoritmos propostos para o roteamento baseado em QoS não deve ser maior que a dos algoritmos utilizados atualmente.

As métricas utilizadas devem refletir as características da rede, de forma a suportar requisitos de QoS básicos. Os parâmetros de QoS negociados deverão sempre ser mapeados na métrica (ou conjunto) usada pelo protocolo de roteamento na decisão de rotas. Desta forma, a escolha da métrica é fator determinante dos tipos de requisitos de QoS que poderão ser suportados. Por exemplo, se o atraso de transmissão e o custo forem as métricas, todos os requisitos terão que ser mapeados em termos de custo e atraso. Neste tipo de métrica, um requisito do tipo probabilidade de perda de pacotes não poderia ser suportado.

As métricas utilizadas devem ser disjuntas, sem qualquer informação redundante. Isto porque a informação redundante pode introduzir inter-dependência entre as métricas, o que torna impossível o cálculo de cada métrica independentemente. O cálculo de rotas torna-se mais complexo com a introdução de algoritmos recursivos entre as métricas.

3.2 Métrica única combinada

Algoritmos para o cálculo de rotas baseados em uma única métrica são bastante conhecidos, sendo utilizados pela maioria dos protocolos de roteamento atuais. A aproximação mais natural seria, então, o mapeamento de diferentes requisitos de QoS do usuário em uma única medida, através da qual a decisão possa ser tomada utilizando algoritmos tradicionais.

Para a definição de uma tal métrica, poderia-se criar uma função geradora, que receba como parâmetros de entrada um conjunto de requisitos de QoS do usuário e forneça na saída uma medida única. Por exemplo, poderia ser definida a função de saída para cada caminho como sendo proporcional à faixa de transmissão disponível e inversamente proporcional ao atraso e à probabilidade de perda de pacotes:

$$f(c) = \frac{B(c)}{D(c) \times L(c)},$$

onde c é um determinado caminho, $B(c)$ é sua banda passante, $D(c)$ o atraso de transmissão associado com este caminho e $L(c)$ a probabilidade de perda de pacotes correspondente. O caminho com o maior valor de $f(c)$ seria o melhor em termos desta medida.

A métrica única combinada, no entanto, não pode ser utilizada na prática, pois não contém informação suficiente para garantir que os requisitos de QoS do usuário estão sendo respeitados. Outro problema advém das diferentes regras de composição para parâmetros de QoS diversos. Para dois caminhos diferentes, a métrica do caminho total será a soma das duas individuais, se esta representar o atraso de transmissão por exemplo. Se, no entanto, a característica de QoS for a banda passante, a métrica do caminho total será a menor das duas. Desta forma, a função definida acima não preserva nenhuma das duas regras de composição.

3.3 Métricas múltiplas

A utilização de métricas múltiplas é capaz de modelar a rede de maneira mais precisa, porém o problema de encontrar um caminho dadas diversas restrições pode se tornar complexo [12, 15]. A quantidade de diferentes requisitos de QoS é um fator de acréscimo da complexidade dos algoritmos. Além disso, os requisitos mais importantes dependem diretamente da aplicação em questão.

A complexidade do algoritmo será dependente da regra de composição das métricas, sendo três as regras de composição mais comuns: a aditiva, a multiplicativa e a côncava. Na aditiva, as métricas dos arcos componentes do caminho devem ser somadas para gerar a métrica do todo. Na multiplicativa, a métrica resultante é o produto das métricas componentes. Na do mínimo, também chamada de côncava, a métrica final é a mínima entre as todas as métricas dos trechos componentes da rota.

Os parâmetros mais comuns utilizados como requisitos de QoS do usuário são o atraso, a variação de atraso (*jitter*), o custo do serviço, a probabilidade de perda de pacotes e a banda passante. O atraso, a variação de atraso e o custo seguem a regra de composição aditiva. A banda passante segue a regra de composição do mínimo. A regra de composição para a probabilidade de perda é um pouco mais complexa. Seja c o caminho cuja métrica será calculada, $d(c)$ a métrica deste caminho, $d(i, j)$ a métrica associada ao arco entre o nó i e o nó j, a regra de composição da probabilidade de perda será como a seguir:

$$d(c) = 1 - ((1 - d(i, j) \times (1 - d(j, k)) \times ... \times (1 - d(l, m))).$$

No entanto, se a probabilidade de perda for transformada na probabilidade de transmissão com sucesso, a métrica equivalente possui uma regra de composição multiplicativa. Em [12] é realizado um estudo acerca da complexidade dos algoritmos para n métricas de forma aditiva e de forma multiplicativa, mostrando que estes são NP-completos. Além disso, é demonstrado que a combinação de n métricas aditivas com m métricas multiplicativas também só é solúvel por algoritmos NP-completos.

Como conseqüência, pode-se concluir que dos requisitos de QoS citados anteriormente, atraso, *jitter*, custo, probabilidade de perda de pacotes e banda passante, as únicas combinações tratáveis seriam da banda passante com outra das métricas. Isto porque a banda passante possui uma regra de composição côncava, enquanto que as outras são aditivas ou multiplicativas.

3.4 Roteamento pela origem e salto-a-salto com duas restrições

A seguir, são apresentados algoritmos para o cálculo de rotas dadas duas restrições, uma delas a faixa de transmissão, a outra podendo ser o atraso, o custo ou o *jitter*, uma vez que todas possuem regras de composição similares.

Na análise a seguir, a banda passante considerada deve ser entendida como a taxa de transmissão residual disponível para os novos fluxos em cada enlace. A métrica banda passante é então definida como a menor banda passante disponível em todos os arcos componentes de um determinado caminho.

O atraso possui dois componentes básicos, o atraso nas filas e o atraso de propagação propriamente dito. É importante notar que o atraso nas filas é função da menor banda passante e das características do tráfego. Desta forma, o atraso utilizado como métrica deve considerar apenas a componente devida ao tempo de propagação, uma vez que a correspondente ao tempo gasto nas filas já está considerada na métrica faixa de transmissão garantindo-se a independência entre as métricas combinadas.

A banda passante mínima de um determinado caminho pode ser entendida como a sua largura, enquanto que o atraso pode ser tomado como o comprimento do caminho. Desta forma, o problema do algoritmo de roteamento passa a ser encontrar um caminho na rede mais curto que determinada restrição, e com uma taxa de transmissão com um certo valor mínimo.

Para muitas aplicações, em especial as de tempo real, o atraso fim-a-fim é um dos requisitos de QoS mais importantes. A utilização da banda passante e do atraso de propagação como métricas constitui um compromisso entre a complexidade e a melhor caracterização dos requisitos de QoS do usuário.

3.4.1 Algoritmo para roteamento pela origem

Nesta seção, é apresentado um algoritmo para o esquema de roteamento pela origem com restrições de faixa de transmissão mínima e de atraso máximo admitido. Seja $G = (N, A)$ um grafo representante da rede, com N nós e E enlaces. A cada arco (i, j) são associados dois números reais, b_{ij} como a banda passante disponível e a_{ij} como o atraso de propagação. Para um arco (i, j) inexistente no grafo, será considerado que $b_{ij} = 0$ e $a_{ij} = \infty$. Dado um caminho $c = (i, j, k, ..., l, m)$, a largura do caminho, $w(c)$, será dada por $min[b_{ij}, b_{jk}, ..., b_{lm}]$ e seu comprimento será a soma dos atrasos de propagação, $l(p) = a_{ij} + a_{jk} + ... + a_{lm}$. Dadas duas constantes B para a faixa de transmissão e A para o atraso máximo aceitável, o problema se reduz a encontrar um caminho c entre i e m tal que $w(c) \geq B$ e $l(c) \leq A$.

O algoritmo consiste primeiro em descartar todos os arcos cuja $b_{ij} < B$. Todos os caminhos possíveis dentro do grafo resultante estarão respeitando o requisito de banda passante. O algoritmo deve então apontar um caminho que respeite o atraso máximo permitido, se ele existir. O algoritmo a seguir encontra um caminho entre o nó 1 e o nó m com banda passante não menor que B e atraso não maior que A. A_i é considerado o comprimento estimado do caminho entre o nó 1 e o nó i.

Passo 1: $\forall i, j : a_{ij} = \infty$ se $b_{ij} = 0$.
Passo 2: $L = \{1\}; \forall i, i \neq 1 : A_i = a_{1i}$.
Passo 3: Encontre $k \notin L \mid A_k = \min_{i \notin L} A_i$.
Se $A_k > A$, o caminho não pode ser encontrado e o algoritmo termina.
Se L contém o nó m, um caminho foi encontrado e o algoritmo termina.
$L := L \cup \{k\}$.
Passo 4: $\forall i \notin L : A_i := \min[A_i, A_k + a_{ki}]$.
Vá para o passo 3.

O passo 1 elimina os arcos cuja faixa de transmissão é menor que B. Os passos 2-4 encontram o caminho de atraso mínimo para o nó m através do algoritmo de Dijkstra. O algoritmo termina se o nó m for incluído no caminho ou se o atraso ultrapassar o limite antes de alcançar o nó m. Cada passo no algoritmo requer da ordem de N operações, e no pior caso os passos são repetidos $N - 1$ vezes. Sendo assim o número de operações total no pior caso é $O(N^2)$.

3.4.2 Algoritmo para roteamento salto-a-salto

Uma vez que o esquema de roteamento salto-a-salto pré-calcula entradas na tabela para todos os destinos possíveis, este deve suportar todos os requisitos de recursos possíveis. Usualmente, os algoritmos de roteamento salto-a-salto calculam o melhor caminho para todos os destinos. Com uma única métrica, a melhor rota pode ser encontrada facilmente. Com mais de uma métrica, no entanto, o melhor caminho, com todos os seus parâmetros possuindo valores ótimos pode sequer existir. Sendo assim, é necessária a definição de uma prioridade entre as diferentes métricas.

A precedência entre a banda passante e o atraso é de certa forma dependente da aplicação. De uma maneira geral, entretanto, o atraso nas filas é mais dinâmico e sensível ao tráfego, sendo a banda passante freqüentemente mais crítica para a maioria das aplicações multimídias. No caso de não haver banda passante suficiente, o atraso nas filas, e provavelmente a taxa de perdas, serão altos. Em contraste, se o atraso de propagação não puder ser satisfeito, o atraso total será maior, mas o acréscimo será previsível e estável. Sendo assim, falhar em qualquer dos dois requisitos aumentará o atraso total do sistema, no entanto as conseqüências da falta de banda passante podem ser mais graves. Sendo assim, nos dois algoritmos a seguir é dada precedência à faixa de transmissão em detrimento do atraso de propagação.

Seja o nó 1 a fonte da transmissão e h o número de arcos saindo do nó fonte. Sejam $B_i^{(h)}$ e $A_i^{(h)}$ a largura e o comprimento do caminho mais curto e mais largo escolhido entre o nó 1 e o nó i com h saltos. Por convenção, $B_1^{(h)} = \infty$ e $A_1^{(h)} = 0$ para todo h. Os algoritmos para descoberta do caminho mais largo e mais curto podem ser produzidos pela adição de um teste de menor comprimento quando forem encontrados caminhos com igual largura. O algoritmo baseado em vetores distância é como a seguir:

Passo 1: Início: $\forall i \neq 1 : h = 0$ e $B_i^{(0)} = 0$

Passo 2: Encontre $K = \{k \mid w(1, ..., k, i) = \max_{1 \leq j \leq N}[\min[B_j^{(h)}, bji]]\}$, $i \neq 1$.

Passo 3: Se $\#K > 1$, encontre $k \in K \mid l(1, ..., k, i) = \min_{1 \leq j \leq N}[A_j^{(h)} + a_{ji}]$, $i \neq 1$.

Passo 4: $B_i^{(h+1)} = w(1, ..., k, i)$ e $A_i^{(h+1)} = l(1, ..., k, i)$.

Passo 5: Se $h \geq A$, o algoritmo termina.

Se não, $h = h + 1$. Vá para o passo 2.

O passo 2 encontra o caminho de maior faixa de transmissão do nó 1 para cada nó i. Se mais de um caminho com a maior faixa for encontrado, o passo 3 elege o caminho com o menor comprimento (atraso). O passo 4 atualiza a largura e o comprimento para o caminho mais largo e mais curto do nó 1 para cada nó i.

Seja o nó 1 a fonte de dados, B_i e A_i a largura e o comprimento do caminho mais curto e mais largo do nó 1 para o nó i. Por convenção, $B_1 = \infty$ e $A_1 = 0$. O algoritmo baseado no estado dos enlaces é o seguinte:

```
Passo 1:  Início:  ∀i ≠ 1 : L = {1}, B_i = b_1i  e  A_i = a_1i.
Passo 2:  Encontre K = {k | k ∈ L, B_k = max_{i∉L}[B_i]}.
Passo 3:  Se #K > 1,  encontre k ∈ K | l(1,...,k,i) = min_{j∈k}[A_(1,...,j,i)].
L := L ∪ {k}.  Se L contém todos os nós, o algoritmo termina.
Passo 4:  ∀i ∉ L,  B_i := max[B_i, min[B_k, b_ki]].
Vá para o passo 2.
```

O passo 2 encontra os nós com a faixa de transmissão máxima entre os nós provisoriamente marcados. Se mais de um nó for encontrado, o passo 3 elege o nó mais curto e o marca permanentemente. O passo 4 atualiza os nós marcados provisoriamente em torno do novo nó permanentemente marcado.

Os dois algoritmos para o cálculo do caminho mais curto e mais largo são escaláveis. Nas duas versões o número de operações necessárias em cada iteração é proporcional às versões correspondentes do algoritmos para o cálculo do caminho mais curto. Sendo assim, o tempo de processamento dos dois algoritmos é igual ao dos algoritmos de caminho mais curto tradicionais.

4 Algoritmos para Fluxos com Três Restrições

Na seção anterior, foram apresentadas as principais dificuldades encontradas na definição de um protocolo de roteamento com requisitos de QoS. A utilização de uma métrica única combinada foi descartada, devido às dificuldades impostas pelas diferentes regras de composição intrínsecas de cada métrica. Em seguida, dadas as considerações sobre a independência entre as métricas [12], foram apresentados três algoritmos. Estes consideram sempre duas métricas, uma delas a faixa de transmissão, que possui uma regra de composição côncava. A outra métrica pode ser o atraso, o *jitter*, ou a probabilidade de perda.

Outro aspecto importante quanto à escalabilidade do protocolo de roteamento diz respeito ao tamanho das tabelas de roteamento. No roteamento baseado em requisitos de QoS, a decisão de rotas é realizada de acordo com as características de cada fluxo. Uma soulução natural seria a definição de tabelas de roteamento com as rotas utilizadas por cada fluxo. A utilização de um tal esquema, no entanto, levaria à construção e manutenção de tabelas muito grandes, devido à quantidade de nós de cada sub-rede e à quantidade de usuários conectados a cada nó.

Uma solução de compromisso entre a precisão da rota utilizada pelo fluxo de acordo com seus requisitos de QoS e o tamanho das tabelas pode ser a manutenção nos roteadores de rotas possíveis para um destino. Estas rotas podem ser divididas em grupos de acordo com a classe de requisitos de QoS do tráfego transportado. O tamanho das tabelas de roteamento é certamente maior que o utilizado pelos protocolos atuais, porém não seria tão grande quanto a manutenção de uma rota por fluxo.

Considerando as observações realizadas, um esquema híbrido pode ser proposto. Este constitui um compromisso entre a simplicidade do algoritmo utilizado e a precisão da definição da rota, de acordo com os requisitos de QoS de um determinado fluxo. Algumas modificações são realizadas nos algoritmos apresentados na seção anterior, através da utilização cuidadosa de uma métrica única combinada. Esta leva em consideração, como exemplo, o atraso de propagação e a probabilidade de perda nos enlaces.

4.1 Combinação de métricas

A utilização da métrica única combinada é uma alternativa interessante uma vez que os algoritmos de roteamento já implementados poderiam ser utilizados sem maiores modificações. No entanto, existem dois empecilhos importantes. O primeiro deles diz respeito à definição de uma função de combinação das métricas compatível com as regras de composição de cada uma. Outro problema é, dada uma métrica que é a combinação de duas ou mais métricas, poder-se garantir que uma determinada rota está garantindo os requisitos de QoS para cada uma das métricas. Em outras palavras, no caso geral, a escolha de um caminho que respeite um limite para a métrica combinada não garante que este caminho respeite cada uma das métricas isoladamente.

As principais métricas consideradas até então são a faixa de transmissão, o atraso de propagação, a variação de atraso e a probabilidade de perda. A faixa de transmissão possui uma regra de composição côncava. O atraso de propagação e o *jitter* possuem ambos regra de composição aditiva (o custo dos enlaces também). Já a probabilidade de perda possui uma regra de composição mais complexa, como mostrado na seção 3.

No entanto, a probabilidade de perda pode ser considerada de outra forma, como sendo a probabilidade de transmissão com sucesso. Neste caso, a regra de composição passa a ser multiplicativa. Seja c o caminho cuja métrica será calculada, $S(c)$ a probabilidade de transmissão com sucesso neste caminho, $S(i,j)$ a métrica associada ao arco entre o nó i e o nó j, a regra de composição da probabilidade de transmissão com sucesso será como a seguir:

$$S(c) = (S(i,j) \times S(j,k) \times ... \times S(l,m)).$$

Se for tomado o logaritmo desta métrica, sua regra de composição passa a ser aditiva. Desta forma, a combinação desta métrica com outras métricas aditivas passa a ser viável. Pode-se definir a probabilidade de transmissão com sucesso logarítmica:

$$Slog(c) = log(S(c)) = (log(S(i,j)) + log(S(j,k)) + ... + log(S(l,m))).$$

É importante notar os efeitos que a transformação logarítmica acarreta. A composição aditiva é desejável, mas outro aspecto deve ser observado. O valor desta métrica, por se tratar de uma probabilidade, estará sempre entre 0 e 1. Quando é tomado o logaritmo, este intervalo é transformado no intervalo de $-\infty$ a 0.

Números com valor absoluto tão grande podem constituir um problema para o cálculo das métricas. Mas se o intervalo considerado para a probabilidade de transmissão com sucesso for reduzido, para entre 0.1 e 1 por exemplo, o intervalo de transformação pode ser restrito, ficando entre -1 e 0. Esta restrição é bastante razoável, uma vez que dificilmente a probabilidade de transmissão com sucesso de um enlace será tão baixa, da ordem de 10%. Na prática, pode-se supor que o valor absoluto da probabilidade logarítmica estará sempre entre 0 e 1. A mesma faixa pode ser esperada para a probabilidade de transmissão com sucesso de um caminho composto de diversos enlaces. Esta probabilidade será sempre menor que a menor das métricas dos enlaces envolvidos, ficando mais próxima de zero.

4.2 Métrica para atraso e probabilidade de perda

Dadas as restrições acima, podemos definir uma métrica única combinada para a probabilidade de perda e o atraso de transmissão. A métrica consiste da soma do atraso de transmissão, dado em microssegundos, somado à probabilidade de transmissão com sucesso logarítmica. Tomando a unidade de atraso de transmissão como o microssegundo, pode-se restringir esta medida a números inteiros. Esta medida é compatível com os atrasos esperados na Internet.

A métrica única combinada é a soma da probabilidade de transmissão com sucesso logarítimica com o atraso de transmissão. A regra de composição das duas métricas é aditiva, não causando problemas. A primeira métrica será sempre representada por números reais entre 0 e 1, e a outra sempre por números inteiros. Desta forma, na métrica total de um caminho poder-se-á sempre distingüir a parte devida a uma e a outra métrica. Esta é outra das desvantagens que eram impostas à utilização de uma métrica única combinada, podendo desta forma ser contornada.

Com a definição de uma métrica única combinada para o atraso e a probabilidade de perda, pode-se, através de um esquema semelhante ao utilizado anteriormente, construir algoritmos que levem em conta também a faixa de transmissão envolvida. A idéia é a mesma, ou seja, realizar uma busca prévia no grafo representativo da sub-rede eliminando os enlaces que não possam suportar a faixa de transmissão requisitada. Desta forma, estes arcos são eliminados do grafo.

4.3 Algoritmo modificado

Nesta seção, são apresentadas pequenas modificações no algoritmo para roteamento pela origem mostrado anteriormente, para levar em consideração a utilização da métrica única combinada definida para o atraso de transmissão e a probabilidade de perda.

Considere-se em primeiro lugar o algoritmo para roteamento pela origem. Seja $G = (N, A)$ um grafo representante da rede, com N nós e E enlaces. A cada arco (i, j) são associados dois números reais, b_{ij} como a banda passante disponível e u_{ij} como a métrica única combinada resultante da soma do atraso de propagação, um inteiro, com a probabilidade de transmissão com sucesso logarítmica, um real entre 0 e 1. Para um arco (i, j) inexistente no grafo, será considerado que $b_{ij} = 0$ e $u_{ij} = \infty$. Dado um caminho $c = (i, j, k, ..., l, m)$, a largura do caminho, $w(c)$, será dada por $min[b_{ij}, b_{jk}, ..., b_{lm}]$ e seu comprimento será a soma das métricas de cada um dos enlaces, $l(p) = a_{ij} + a_{jk} + ... + a_{lm}$. Dadas duas constantes B para a faixa de transmissão e U para o valor máximo aceitável da métrica única, o problema se reduz a encontrar um caminho c entre i e m tal que $w(c) \geq B$ e $l(c) \leq U$, como originalmente.

Passo 1: $\forall i, j : u_{ij} = \infty$ se $b_{ij} = 0$.
Passo 2: $L = \{1\}; \forall i, i \neq 1 : U_i = u_{1i}$.
Passo 3: Encontre $k \notin L \mid U_k = \min_{i \notin L} U_i$.
Se $U_k > U$, o caminho não pode ser encontrado e o algoritmo termina.
Se L contém o nó m, um caminho foi encontrado e o algoritmo termina.
$L := L \cup \{k\}$.
Passo 4: $\forall i \notin L : U_i := \min[U_i, U_k + u_{ki}]$.
Vá para o passo 3.

O algoritmo como mostrado acima não garante que a probabilidade de transmissão com sucesso será maior que a desejada (logaritmicamente, menor). No entanto, como o atraso está representado por um número inteiro, este estará sempre respeitado. No entanto, dados dois enlaces com um mesmo atraso, está garantido que a probabilidade de perda do enlace escolhido será sempre a menor.

Uma pequena modificação pode ser incluída no algoritmo, de forma que a probabilidade de transmissão com sucesso seja garantida, de uma maneira aproximada. O requisito de probabilidade de perda se traduz em um valor (absoluto) mínimo para o logaritmo da probabilidade de transmissão com sucesso. Ou seja, a soma desta métrica para todos os enlaces de um caminho deve ser menor que um limite.

Seja P o limite para o logaritmo da probabilidade de transmissão com sucesso, e $p(ij)$ o logaritmo da probabilidade de transmissão no enlace ij. Dado um caminho c com x saltos, se $p(ij)$ para cada caminho for menor que P/x, $p(c)$ será menor que P. No entanto, no momento em que um caminho está sendo calculado, o seu tamanho não é conhecido. Uma idéia seria restringir $p(ij)$ a P/C, onde C é o tamanho do maior caminho no grafo representativo da sub-rede. Porém este limite pode ser muito restritivo. O caminho entre dois nós será em geral menor que C, sendo assim sua probabilidade de perda estará mais restrita que o necessário.

Um esquema alternativo pode relaxar um pouco esta restrição. A cada passo do algoritmo, define-se um "saldo" de saltos (SS) e um "saldo" de probabilidade logarítmica (SP). O saldo de saltos consiste na diferença entre o tamanho do maior caminho e o número de saltos do caminho no passo atual do algoritmo. O saldo de probabilidade é a diferença entre o limite exigido e a probabilidade logarítmica total do caminho construído. A probabilidade logarítmica máxima para o próximo enlace será não mais P/C, mas SP/SS. Este valor pode ser maior que P/C, caso o SP seja menor que $SP \times (P/C)$.

Desta forma, o algoritmo deve ser modificado:

```
Passo 1:   ∀i,j : u_ij = ∞ se b_ij = 0.
SS = C; SP = P.
Passo 2:   L = {1}; ∀i, i ≠ 1 : U_i = u_1i.
Passo 3:   Encontre k ∉ L | (U_k = min_{i∉L} U_i) e ((U_k − ⌊U_k⌋) < (SP/SS)).
Se U_k > U , o caminho não pode ser encontrado e o algoritmo termina.
Se L contém o nó m, um caminho foi encontrado e o algoritmo termina.
L := L ∪ {k}.
SS := SS − 1.
SP = SP + (U_k − ⌊U_k⌋).
Passo 4:   ∀i ∉ L : U_i := min[U_i, U_k + u_ki].
Vá para o passo 3.
```

Desta forma, a restrição quanto à probabilidade de transmissão com sucesso continua sendo mais restrita, para caminhos menores que o caminho máximo, que a restrição real. Uma rota possível para determinado fluxo pode ser ignorada pelo algoritmo, como no caso anterior. Esta restrição é no entanto um pouco mais relaxada, e garante que ambas as métricas serão respeitadas sem aumentar a complexidade do algoritmo. Esta imprecisão na descoberta de rotas constitui um compromisso com a complexidade do algoritmo utilizado. As mesmas modificações podem ser realizadas nos outros dois algoritmos apresentados na seção anterior, levando à descoberta de rotas dadas três restrições, a faixa de transmissão mínima, o atraso e a probabilidade de perda.

5 Conclusões

O roteamento baseado em requisitos de QoS deve ser considerado como um componente fundamental na provisão de novas classes de serviço na Internet, que possam dar suporte à variedade de novas aplicações multimídias que vêm surgindo. Este esquema possui diversos pré-requisitos a serem respeitados, sendo que cada um de seus componentes, como o cálculo de rotas, a escalabilidade e o controle administrativo, apresenta questões específicas a serem consideradas.

Os algoritmos para cálculo de rotas são um dos principais focos deste esquema de roteamento. Existem duas soluções possíveis para a realização de descoberta de rota dadas várias métricas. Estas consistem em definir uma métrica única combinada, ou em tratar as diversas métricas separadamente. A primeira solução, embora possua características que a inviabilizam, teria como grande vantagem a utilização dos algoritmos já definidos para os protocolos de roteamento atuais.

Alguns autores consideram a utilização de algoritmos considerando mais que duas restrições inviável, dada sua complexidade. No caso de uma caracterização de tráfego através de duas restrições, algoritmos simples podem ser construídos [12]. Através da combinação destes algoritmos com uma métrica única combinada de outras duas, podem ser definidos algoritmos para a construção de rotas dadas três restrições. Neste trabalho, foi realizada a proposta de algoritmos simples que levam em conta como restrições a faixa de transmissão, a probabilidade de erro e o atraso de transmissão. Estes possuem quase a mesma complexidade que os algoritmos propostos anteriormente para tráfegos com duas restrições. No algoritmo proposto, uma métrica de composição multiplicativa pôde ser combinada com uma de composição aditiva.

A definição de algoritmos simples é importante para a implementação de protocolos de roteamento baseados em requisitos de QoS escaláveis. Para grande parte das aplicações, a utilização de três restrições é suficiente para caracterizar as necessidades de QoS.

A definição de protocolos de roteamento baseados em requisitos de QoS é, portanto, uma questão importante na definição de uma arquitetura de Serviços Integrados para a Internet. Algumas extensões aos protocolos de roteamento existentes estão sendo propostas, no entanto não foi definido qualquer novo protocolo. A definição de protocolos baseados nos algoritmos apresentados neste trabalho mostra-se uma área promissora para pesquisas futuras.

Referências

[1] R. Braden, D. Clark e S. Shenker, *RFC1633: Integrated Services in the Internet Architecture: An Overview.* Network Working Group, junho 1994.

[2] L. Zhang, S. Deering, D. Estrin, S. Shenker e D. Zappala, "RSVP: A new resource reservation protocol", *IEEE Network Magazine*, setembro 1993.

[3] L. Delgrossi, R. G. Herrtwichi, C. Vogt e L. C. Wolf, "Reservation protocols for internetworks: A comparison of ST-II and RSVP", in *Proceedings of Fourth International Workshop on Network and Operating System Support for Digital Audio and Video*, Lancaster, UK, novembro 1993.

[4] D. J. Mitzel, D. Estrin, S. Shenker e L. Zhang, "An architectural comparison of ST-II and RSVP", in *Proceedings of IEEE Infocom*, 1994.

[5] F. S. G. Oliveira, "Mobilidade na internet: Mecanismos básicos, otimizações e desafios", tech. rep., PEE-COPPE/UFRJ, 1997.

[6] B. Rajagopalan e R. Nair, *Quality of Service (QoS)-Based Routing in the Internet - Some Issues.* Bellcore and Ascon Nexen, outubro 1996. Internet-draft.

[7] E. Crawley, R. Nair, B. Rajagopalan e H. Sandick, *A Framework for QoS-Based Routing in the Internet.* Gigapacket Networks, Arrowpoint, NEC USA e IBM, julho 1997. Internet-draft.

[8] M. Steenstrup, *RFC1477: IDPR as a proposed standard.* Network Working Group, julho 1993.

[9] S. Berson e L. Berger, *IP Integrated Services with RSVP over ATM.* ISI and Fore Systems, março 1997. Internet-draft.

[10] D. J. Mitzel, D. Estrin, S. Shenker e L. Zhang, “A study of reservation dynamics in integrated services packet networks”, *IEEE Network Magazine*, pp. 18–26, jan-fev 1995.

[11] H. G. Perros e K. M. Elsayed, “Call admission control schemes: a review”, *IEEE Communications Magazine*, novembro 1996.

[12] Z. Wang e J. Crowcroft, “Quality of service routing for supporting multimedia applications”, *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 7, pp. 1228–1234, setembro 1996.

[13] J. Moy, *RFC1584: Multicast Extensions to OSPF.* Network Working Group, março 1994.

[14] J. Moy, *RFC1585: MOSPF: Analysis and Experience.* Network Working Group, março 1994.

[15] Q. Ma e P. Steenkiste, “Quality of service routing for traffic with performance guarantees”, in *IFIP Fifth International Workshop on Quality of Service*, New York, pp. 115–126, maio 1997.