Título do artigo: Arquitetura de Integração de Serviços para a Comutação IP

Responsável: José F. de Rezende
E-mail: rezende@gta.ufrj.br

Arquitetura de Integração de Serviços para a Comutação IP

José F. de Rezende[(1)] e Simon Pietro Romano[(2)]

[(1)]Coppe Elétrica
Universidade Federal do Rio de Janeiro
rezende@gta.ufrj.br

[(2)]Gruppo di Ricerca Informatica Distribuita
Universitá di Napoli Federico II - Naples - Italy
sprom@grid.grid.unina.it

Resumo

Neste artigo nós investigamos os aspectos relacionados à integração da QoS na tecnologia de comutação IP. Primeiramente, nós propomos uma arquitetura que inclui o protocolo RSVP com um mínimo de impacto na estrutura original e nos protocolos da comutação IP. Isso nos levou a um estudo mais detalhado dos mecanismos de escalonamento e de filas adotados nos níveis de roteamento e de comutação, de forma a garantir um suporte efetivo aos serviços Garantido (GS) e Best-Effort (BE), como definido pelo IETF. A partir da observação de que o serviço GS resulta normalmente numa má utilização dos recursos da rede, nós então introduzimos uma nova classe de serviço (BR). Esta classe provê garantias aos fluxos de dados em termos da taxa em longo prazo, sem afetar os fluxos GS e explorando a banda passante não utilizada pelos mesmos. Finalmente, propomos uma implementação deste novo serviço e a validamos com a ajuda de simulação.

Abstract

In this paper we investigate issues related to the integration of QoS capabilities in the framework of IP switching technology. First, we propose an architecture which includes the RSVP protocol with minimal impact on the original structure and protocols of IP switching. This leads us to further investigate the queueing and scheduling mechanisms adopted both at the routing and the switching level in order to guarantee effective support for Guaranteed (GS) and Best-Effort (BE) services, as defined by the IETF. Based on the observation that GS service typically results in network underutilization, we then introduce a new service class (BR) which provides flows with long-term rate guarantees without impacting GS flows and while exploiting their unused bandwidth. Finally, we propose a possible implementation of this new service and we validate it with the help of simulation.

1 Introdução

Novas aplicações, tais como a telemedicina, o trabalho cooperativo suportado por computador (CSCW) e os sistemas de simulação distribuída, são bastante exigentes quanto à QoS recebida pelo serviço de comunicação. Com o objetivo de fornecer a estas aplicações um suporte adequado, a comunidade Internet, através do grupo de trabalho *Integrated Services* (IntServ) do IETF, especificou um modelo de trabalho para o controle da QoS [1, 2, 3] oferecida pelo protocolo IP. Por outro lado, o ATM Forum definiu para as redes ATM várias categorias de serviço apropriadas às diversas naturezas de tráfego gerado por essas aplicações [4].

O protocolo IP vem exercendo há anos um papel primordial na interconexão global de redes de computadores. A natureza orientada sem conexão deste protocolo tem provado ser um aspecto chave no suporte de redes de integração de serviços. Por outro lado, a instalação ubíqua das redes ATM levanta muitas questões importantes relacionadas à interoperação das arquiteturas IP e ATM. Os diferentes paradigmas adotados por tais redes (orientado sem conexão e orientado conexão) fazem da sua coexistência uma tarefa difícil. Uma outra preocupação é a integração da QoS num ambiente IP/ATM.

Diferentes propostas foram apresentadas visando resolver o problema de IP sobre ATM. Uma das soluções que consideramos promissora é a comutação IP, como apresentada por Ipsilon [5]. A comutação IP se baseia na completa eliminação da sobrecarga imposta pela sinalização ATM, e assim preservando a natureza orientada sem conexão do protocolo IP. Nesta solução, uma parte dos fluxos é comutada (ou chaveada) pelo elemento comutador ATM (*IP switch*) no nível 2, ao invés de serem enviados ao nível 3 para serem encaminhados pelos roteadores IP. A comutação de fluxos IP em hardware elimina os atuais gargalos de desempenho impostos pelo motor de encaminhamento nó por nó.

No entanto, a qualidade de serviço recebida por cada fluxo IP atravessando um comutador IP depende de um processo de classificação do fluxo feito de acordo com políticas locais ao comutador. Na atual implementação, tal processo decide se um fluxo deve ser comutado de acordo com o tipo de aplicação que gera o fluxo [5]. Portanto, a QoS obtida por cada fluxo IP depende da distribuição do tráfego na rede Internet e das potencialidades do hardware ATM subjacente, ou seja, das capacidades de escalonamento oferecidas pelo elemento comutador. Para possibilitar uma qualidade de serviço individual para cada fluxo IP, um mecanismo de sinalização deve ser fornecido a fim de permitir que uma aplicação realize pedidos de reserva de recursos. Além disso, os mecanismos de escalonamento do comutador ATM devem ser configurados de forma a reservar tais recursos. RSVP é um protocolo de sinalização que vem recebendo a máxima atenção [6]. Ele é um protocolo orientado receptor e usa um *soft state* para reservar recursos nos elementos da rede. Neste contexto, a questão principal é o processo através do qual os fluxos IP são selecionados para serem comutados, levando-se em conta a reserva de recursos solicitada. Em outras palavras, o processo de classificação de fluxos baseado na QoS. Um outro ponto importante diz respeito ao controle de forma integrada da QoS do tráfego total (roteado e comutado) que atravessa um comutador IP.

Neste artigo nós investigamos como a comutação IP pode prover uma QoS individual para os múltiplos fluxos IP e qual suporte esta solução requer do hardware ATM subjacente. Além disso, nós analisamos como o protocolo RSVP pode ser integrado nesta arquitetura e como fornecer um *framework* para controlar a qualidade de serviço de múltiplas classes de tráfego. O artigo é organizado em cinco seções. Os pontos principais da solução Ipsilon são apresentados na seção 2. Esta seção também discute como o RSVP pode ser integrado na comutação IP. A seção 3 introduz uma arquitetura com integração de serviços para a comutação IP. Nesta seção, nós apresentamos uma nova classe de serviço que permite uma melhor utilização dos recursos da rede e discutimos os mecanismos de filas, de escalonamento e de controle de admissão. Na seção 4 discutimos os resultados de simulação obtidos e, finalmente, na seção 5 apresentamos as conclusões finais deste trabalho.

2 Comutação IP e RSVP

A comutação IP, como proposta por Ipsilon, consiste na especificação de dois protocolos: GSMP (*General Switch Management Protocol*) e IFMP (*Ipsilon Flow Management*

Protocol). Um equipamento de comutação IP (*IP switch*) é construído a partir de dois componentes, um elemento comutador ATM (*ATM switch fabric*) e o Controlador do Comutador IP (*IP Switch Controller*). O GSMP [7] permite que o Controlador do Comutador IP tenha um controle total do elemento comutador ATM. Através do GSMP, o Controlador pode manipular a tabela de comutação ATM e, ao mesmo tempo, o elemento comutador ATM pode informar o Controlador sobre as suas próprias potencialidades. Por sua vez, o protocolo IFMP [8] permite a interação entre entidades remotas do Controlador através de cada enlace conectando dois comutadores IP. Ele torna possível a distribuição de informações guardadas em cache usadas no encaminhamento do fluxo. Além disso, este protocolo permite a um nó atribuir uma etiqueta, particularmente o par VPI/VCI, a um fluxo específico. Depois de informar os vizinhos acima e abaixo sobre a etiqueta associada a um determinado fluxo IP, um nó pode comutar pacotes pertencentes a este fluxo no nível 2 ao invés de encaminhá-los ao nível 3 (figura 1). A operação que permite etiquetar um fluxo e trocar esta informação entre entidades IFMP remotas é chamada de redirecionamento de fluxo (*flow redirection*).

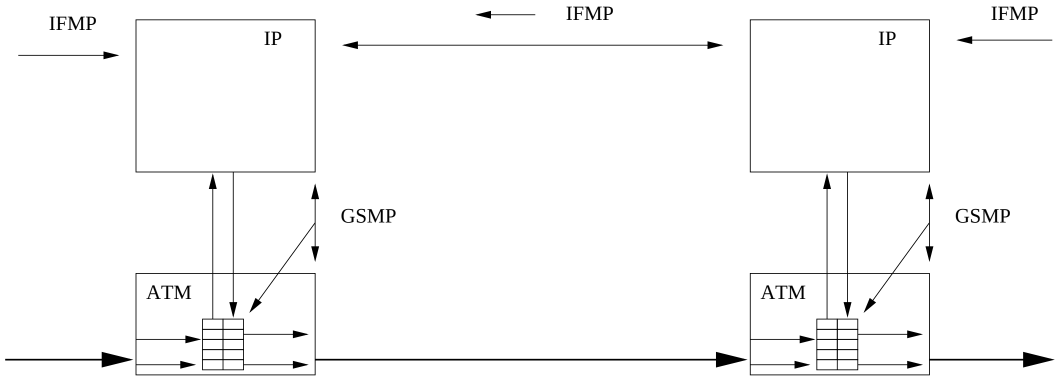


Figura 1: A comutação IP

A decisão de atribuir uma etiqueta a um fluxo IP e assim passar a comutá-lo é tomada pelo módulo de classificação de fluxos de acordo com políticas locais ao comutador IP. Usualmente esta classificação é baseada nas características do fluxo, tais como a aplicação que gera o fluxo, a informação contida no campo *Tipo de Serviço* (TOS), o protocolo utilizado, etc. A cada etiqueta é associado um identificador de fluxo contendo os valores de um conjunto de campos do cabeçalho de um pacote IP, que um pacote deve conter para pertencer a este fluxo. Para manter um *soft state*, tais etiquetas envelhecem com o tempo, ou seja, depois da expiração do seu tempo de vida sem um refrescamento a associação etiqueta-fluxo correspondente é apagada.

Para permitir que as aplicações possam pedir um controle ativo da QoS fim-a-fim ao longo do caminho de um fluxo de pacotes, um protocolo de sinalização deve ser integrado a esta arquitetura. Este protocolo de sinalização servirá então como um *front-end* entre a aplicação que solicita o controle de QoS e o serviço de comunicação. O protocolo RSVP contém as características mais relevantes exigidas pelas aplicações para indicar seus requisitos em termos de QoS. Ele é um protocolo iniciado pelo receptor e, da mesma forma que o protocolo IFMP, utiliza um *soft state* para manter as informações sobre os recursos da rede dedicados a uma determinada comunicação.

Os dados tratados pelo protocolo RSVP são de três naturezas segundo a entidade que fornece (emissor ou receptor) ou modifica (elementos intermediários da rede) tais dados. A informação fornecida por cada emissor e transportada no objeto *Sender TSpec* indica o tipo de tráfego que ele pretende gerar. O objeto *FlowSpec* fornecido por cada receptor é constituído de duas partes, o *Receiver TSpec* e o *RSpec*. O primeiro contém a descrição de tráfego ao qual a reserva de recurso será aplicada, enquanto o segundo carrega a classe de serviço a ser usada e os parâmetros de qualidade de serviço correspondentes. A informação modificada pelos nós intermediários da rede é transportada no objeto ADSPEC da mensagem PATH, o qual contém os serviços disponíveis, as estimativas de retardo e de banda passante, e parâmetros adicionais específicos a um determinado serviço. Esta informação é usada pelos receptores para escolher um serviço e para determinar os parâmetros de reserva de recursos. A especificação do RSVP não define *a priori* o formato interno dos campos do protocolo ou dos objetos citados acima. No entanto, eles são especificados por cada serviço de controle de QoS definido de acordo com as suas necessidades.

Nós argumentamos que os protocolos IFMP e RSVP podem ser de alguma forma integrados para fornecer um controle de QoS individual a cada fluxo IP. Desta maneira, a decisão de comutar um fluxo é totalmente determinada pelas garantias de QoS exigidas pelas aplicações e não por políticas locais adotadas por cada comutador IP. Por estes dois protocolos terem funcionalidades duplicadas, através do *piggybacking* de mensagens de redirecionamento do protocolo IFMP em mensagens RESV do protocolo RSVP, podemos ter um controle mais consistente da QoS fim-a-fim. Na verdade, alegamos que a separação entre a fase de reserva de recursos e a fase de redirecionamento de fluxos pode criar estados incongruentes. Além disso, a natureza *soft state* destes protocolos pode ser uma fonte de problemas na tentativa de associar um tempo de vida à reserva de recursos. E ainda, um aspecto importante é que podemos, através da integração destes protocolos, reduzir a sobrecarga imposta pela execução em separado dos dois protocolos.

Para alcançar a integração IFMP-RSVP propomos adicionar um novo estilo de reserva, chamado *IFMP-style*, ao protocolo RSVP. Esta extensão permitiria o encapsulamento de mensagens de redirecionamento do protocolo IFMP na forma de novos objetos no interior da lista de descritor de fluxo das mensagens RESV. O que deve ser deixado claro, no entanto, é que não estamos fazendo nenhuma alteração no RSVP além do acréscimo do estilo de reserva IFMP. Na recepção de mensagens RESV com este novo estilo, o módulo RSVP passa ao módulo IFMP a informação contida no objeto RESV correspondente. Depois de processar esta informação, o protocolo IFMP retorna ao módulo RSVP a nova mensagem de redirecionamento, a qual deve ser enviada ao vizinho na direção acima. Novamente, esta mensagem é encapsulada em uma nova mensagem RESV. Na chegada de uma mensagem RESV no emissor, todas as associações etiqueta-fluxo já estarão feitas ao longo do caminho a ser percorrido pelo fluxo IP exigindo garantias de QoS. Em outras palavras, graças ao encapsulamento, as mensagens RESV enviadas pelo receptor automaticamente engatilham a criação da associação etiqueta-fluxo.

3 Arquitetura de integração de serviços

Nossa arquitetura de integração de serviços para a comutação IP (figura 2) é organizada em duas partes. A parte relativa ao roteamento consiste basicamente na arquitetura clássica de um roteador IP acrescido de funcionalidades ligadas à QoS. No entanto, algumas melhorias foram feitas com a finalidade de dar suporte às novas potencialidades disponibilizadas pela comutação IP. Nesta arquitetura encontramos um classificador de pacotes, um escalonador de pacotes, um classificador de fluxos e os módulos RSVP, IFMP e GSMP. A parte associada à comutação é constituída pelo comutador ATM, o qual deve prover as funcionalidades relativas à um escalonamento do tipo WFQ. A reserva de recursos no nível de comutação é feita pela configuração do hardware associado ao sistema de filas e ao escalonamento do comutador ATM. Enquanto as funções de policiamento de tráfego são realizadas através da configuração do hardware associado a UPC (*Usage Parameter Control*) de acordo com o parâmetro *flowspec* presente na mensagem RESV.

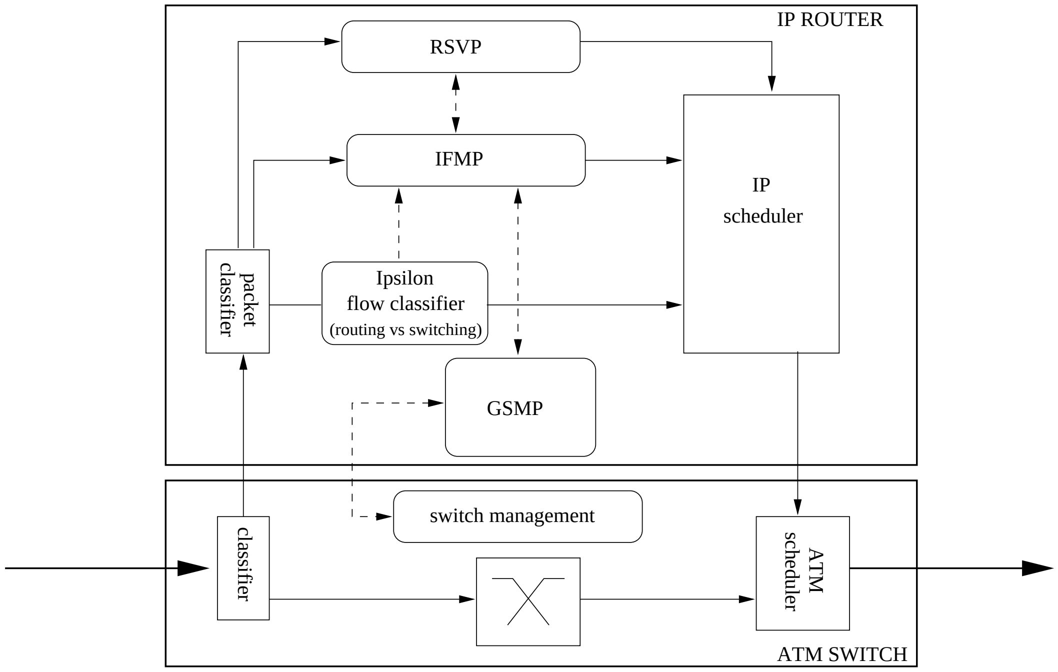


Figura 2: Arquitetura de integração de serviços para a comutação IP.

3.1 Escalonamento no nível roteamento

As políticas de escalonamento no nível roteamento levam em conta dois tipos diferentes de fluxos: i) tráfego de sinalização (i.e. tráfego RSVP/IFMP); e ii) tráfego elástico *best-effort*. O tráfego de sinalização, por ser considerado vital, recebe uma prioridade maior do que aquela associada ao tráfego *best-effort*. O escalonador vem, então, a ser do tipo FIFO priorizada com duas filas, uma para o tráfego de sinalização e a outra para o tráfego *best-effort*.

3.2 Escalonamento no nível comutação

No nível de comutação, todas as classes de serviço são levadas em conta uma vez que nós assumimos que sempre que possível todos os fluxos exigindo QoS serão comutados ao invés de roteados. Reconhecemos que o espaçamento VC pode ser uma fonte de problemas neste caso. Todavia, argumentamos que se por um lado as políticas de tarifação limitarão a demanda de garantias de QoS; por outro lado, a contínua evolução da tecnologia de hardware possibilitará a construção de comutadores cada vez mais possantes. Além disso, técnicas de agregação de fluxos poderão fornecer soluções plausíveis para a superação desta limitação.

3.2.1 Serviço Garantido (GS)

A classe de serviço GS provê limites máximos no retardo de transferência fim-a-fim sofrido por cada pacote do fluxo ao atravessar as filas de espera dos elementos da rede [9]. Além de um limite máximo no retardo de transferência, este serviço garante ao usuário uma certa banda passante. O GS emula o serviço que seria oferecido por um canal de comunicação dedicado entre o emissor e o receptor. Dois parâmetros se aplicam a este serviço: o *TSpec* e o *RSpec*. O primeiro descreve as características de tráfego para o qual o serviço está sendo pedido. Este parâmetro é composto por um descritor de tráfego do tipo *token bucket* mais uma taxa de pico (p), uma unidade de policiamento mínimo (m), e um tamanho máximo de datagrama (M). O *RSpec* especifica a QoS que um dado fluxo pede aos elementos da rede e tem a forma de uma taxa de limpeza (*clearing rate*) R e um termo de folga (*slack term*) S. A taxa de limpeza corresponde a taxa necessária para *limpar* uma rajada de pacotes que pode se armazenar nas filas de espera de um elemento da rede, e garantir assim um limite máximo no retardo de transferência. Por sua vez, o termo de folga indica a diferença entre o retardo de transferência desejado e o garantido depois que o receptor escolheu um valor para R.

Na nossa arquitetura, fluxos do serviço GS são sempre tratados diretamente no nível de enlace de dados. Desta maneira, nós somos capazes de evitar o mecanismo de escalonamento do nível IP obtendo um menor valor para a taxa de limpeza R para um mesmo requisito de limite no retardo de transferência. Esta política leva a uma melhor utilização dos recursos da rede.

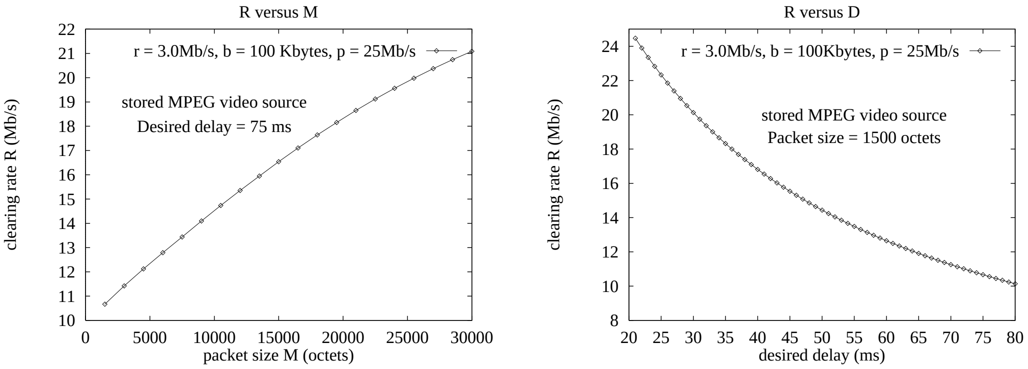
Durante a fase de reserva de recursos, cada elemento intermediário da rede i computa C_i e D_i, os quais representam o desvio de um servidor fluído que opera a uma taxa de limpeza R solicitada por um receptor potencial de um dado fluxo. Estes valores são acumulados nos campos C_{tot} e D_{tot} presentes no objeto ADSPEC de uma mensagem PATH. A equação abaixo [10] aplica-se à computação do limite máximo no retardo fim-a-fim para elementos da rede com tamanho da MTU (Maximum Transfer Unit) igual a L e um fluxo caracterizado por (b, r, p), onde b é igual ao tamanho da rajada, e r e p são as taxas média e de pico de geração de pacotes, respectivamente :

$$\overline{D}_i = \begin{cases} \dfrac{(b-L)}{R}\dfrac{(p-R)}{(p-r)} + \dfrac{L}{R} + \displaystyle\sum_{j=1}^{i}\left[\frac{C_j}{R} + D_j\right] & \text{if } p > R, \\ \dfrac{L}{R} + \displaystyle\sum_{j=1}^{i}\left[\frac{C_j}{R} + D_j\right] & \text{if } p \leq R. \end{cases} \tag{1}$$

O procedimento usado para determinar o valor da taxa de limpeza a ser inserida na mensagem RESV é o seguinte: (i) substituir $R = p$ na segunda das fórmulas acima e computar o valor correspondente do limite superior no retardo de transferência ($\overline{D}$); (ii) se $\overline{D} < D_{desired}$, então inserir os valores conhecidos dos parâmetros de caracterização de fluxo e o limite desejado no retardo de transferência ($D_{desired}$) na primeira fórmula para determinar o valor correspondente da taxa de limpeza R; (iii) caso contrário, utilizar a segunda fórmula.

Como exemplo, considere um enlace OC-3 (155 Mb/s) de um elemento da rede ATM. Assumimos um fluxo de vídeo MPEG armazenado em disco com uma taxa média de aproximadamente 3.0 Mb/s e um tamanho máximo de rajada de 100 Kbytes. A taxa de pico é limitada pela velocidade do meio físico, ao qual a fonte está ligada, isto é, um enlace de 25 Mb/s UTP. Consideramos que cada fluxo atravessa 5 nós com um retardo total de propagação de 20 ms. Subtraindo o retardo de propagação do retardo de transferência desejado obtemos o retardo de transferência causado pelas filas de espera dos nós da rede. Todos os parâmetros descritos acima precisam ser convertidos para levar em conta a sobrecarga adicionada pelo encapsulamento LLC/SNAP e o cabeçalho da célula ATM. Substituindo o retardo de transferência devido às filas de espera em (1), obtemos a taxa que precisa ser reservada para cada fluxo. As figuras 3(a) e 3(b) mostram a taxa de limpeza reservada em função do tamanho máximo do pacote (M) e do retardo de transferência exigido (D), respectivamente. Na figura 3(a), o retardo de transferência exigido, que precisa ser traduzido numa taxa de reserva R, é de 75 ms.

O Serviço Garantido resulta numa fraca utilização da banda passante do enlace, uma vez que a taxa reservada R é sempre maior que a taxa média do fluxo. Uma solução deve ser encontrada para permitir que outros fluxos explorem a banda passante não utilizada. Tais fluxos devem apresentar requisitos menos rigorosos do que os fluxos GS. Na classe



(a) R versus M (b) R versus D

Figura 3: Fonte de vídeo MPEG armazenado.

de serviço descrita abaixo, os fluxos são capazes de recuperar esta banda passante não utilizada, conseguindo alguma garantia da banda passante a longo prazo. A idéia por trás deste serviço é similar àquela apresentada em [10]. A principal diferença reside na maneira em que este serviço é implementado e como ele interage com as demais classes de serviço, tais como o best-effort.

3.2.2 Serviço BR

Muitas das aplicações existentes não necessitam de garantias estritas em termos dos limites no retardo, mas requerem um certo nível de garantia em relação a banda passante. Na verdade, a maioria das aplicações podem se satisfazer com uma taxa mínima de serviço mais um compartilhamento justo da banda passante residual. No serviço best-effort, por exemplo, a taxa mínima de serviço é igual a zero. A introdução de uma classe de serviço que satisfaça estes requisitos seria de grande valia numa arquitetura de serviços integrados. O tráfego gerado pelas aplicações Web é um ótimo exemplo de fluxos que se beneficiariam deste tipo de serviço.

Para implementar esta classe de serviço, chamada Serviço BR (Bandwidth Recovery), usamos um mecanismo de escalonamento WFQ com filas separadas para cada fluxo GS e BR. A cada fluxo GS um peso é atribuído determinando a taxa de limpeza reservada para este fluxo. Assumindo M fluxos GS caracterizados por um envelope (b, r, p) e um enlace com capacidade C_l, para garantir os limites exigidos no retardo de transferência a seguinte desigualdade deve ser satisfeita:

$$\sum_{i=1}^{M} R_i \; \leq \; C_l - \Delta_{\mathrm{BR}} - \Delta_{\mathrm{SIG}} \tag{2}$$

Δ_{BR} e Δ_{SIG} na equação acima representam as porções da banda passante reservadas aos fluxos BR e de sinalização, respectivamente. O peso associado ao fluxo de sinalização é baseado numa estimativa da banda passante necessária para o tráfego correspondente. Enquanto que o peso atribuído a cada fluxo BR é propositadamente muito pequeno e não representa a taxa mínima exigida por cada fluxo. A função do peso é permitir que o fluxo ganhe acesso à banda passante residual deixada pelos fluxos GS sem fornecer nenhuma garantia em termos de limites no retardo. Considerando M' fluxos BR com taxa média igual a r', para garantir suas taxas mínimas a longo prazo a seguinte desigualdade se aplica:

$$\sum_{i=1}^{M} r_i + \sum_{i=1}^{M'} r'_i \leq C_l - \Delta_{\text{SIG}} \tag{3}$$

A equação acima assegura que uma vez que os fluxos GS respeitem seus contratos de tráfego ou um mecanismo de policiamento apropriado (e.g. algoritmo GCRA na UPC) seja utilizado, os fluxos BR receberão uma taxa mínima garantida depois de um certo período de tempo. Na nossa arquitetura, os fluxos BR são sempre comutados no nível de enlace de dados.

3.2.3 Serviço Best-Effort

Os fluxos BE não entram no escopo do nosso estudo, uma vez que eles não exigem nenhuma reserva de recursos e assim não utilizam o protocolo RSVP. Todavia, eles são integrados na nossa arquitetura de QoS como fluxos de baixa prioridade no mecanismo de escalonamento WFQ no nível de enlace de dados. Todos os fluxos BE são misturados em uma única fila do escalonador e são somente servidos quando nenhum outro tráfego está presente nas filas de mais alta prioridade. Isto evita que os fluxos BE prejudiquem fluxos exigindo QoS. O tráfego *best-effort* é roteado ou comutado de acordo com as políticas propostas por Ipsilon [5] e outros [11].

A figura 4 mostra o mecanismo de escalonamento adotado no nível de comutação da nossa arquitetura. Nesta figura, representamos todos os fluxos, isto é, os fluxos, GS, BR e BE, com as suas respectivas filas.

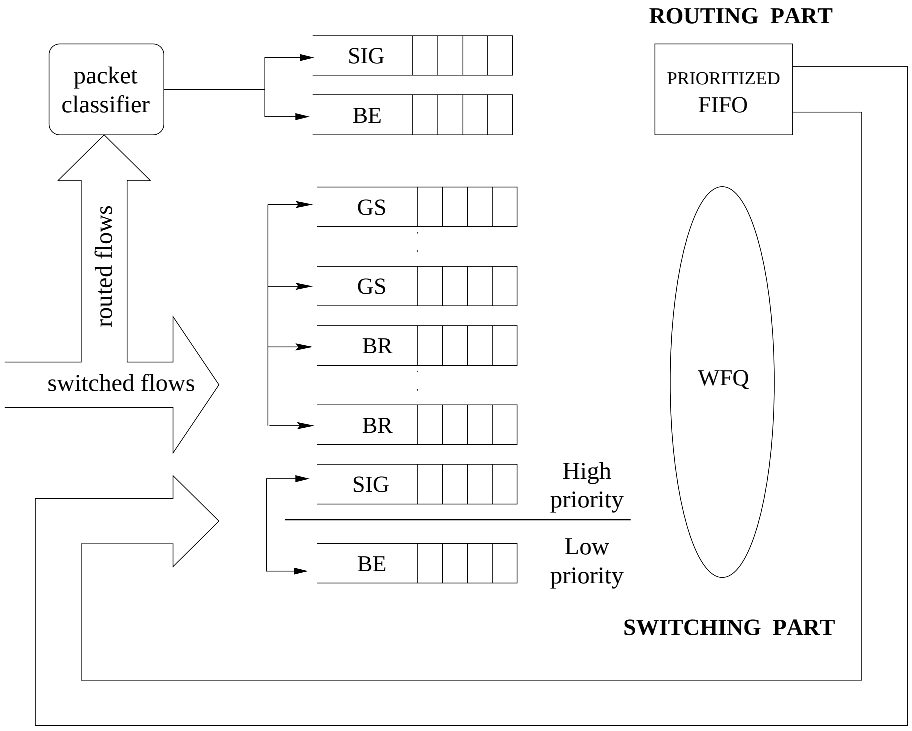


Figura 4: Mecanismo de escalonamento.

4 Resultados de simulação

Em nossas simulações, utilizamos uma rede (figura 5) com 5 comutadores IP. Cada comutador tem no nível de comutação um escalonador WFQ com filas para cada fluxo GS e

BR e uma fila para todos os fluxos BE. A cada comutador IP é conectada uma estação que gera e recebe vários fluxos. Um total de 9 fluxos GS são distribuídos de tal forma que cada enlace inter-comutador seja compartilhado por 6 fluxos GS. Existem 3 fluxos que atravessam 5 enlaces, 2 fluxos que atravessam 4, 2 fluxos que atravessam 3 e 2 fluxos que atravessam 2.

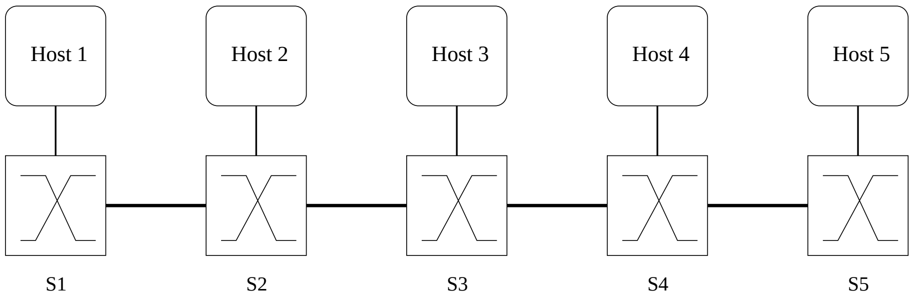


Figura 5: Modelo de simulação da rede.

Todas as fontes de tráfego são modeladas através de uma cadeia de markov com dois estados. Para simular um tráfego de tempo real, em cada período de rajada (ON) um número de pacotes aleatório com distribuição geométrica é gerado a uma taxa de pico p; $\overline{b}$ indica o tamanho médio de cada rajada. O período de silêncio (OFF) tem duração média I. A taxa de geração média de pacotes é dada por $r = \rho \times p$, onde ρ indica a probabilidade que o processo de markov se encontre no estado ON. Dada as relações: $\alpha = (\text{MEANON})^{-1} = \frac{p}{\overline{b}}$, $\beta = (\text{MEANOFF})^{-1} = \frac{1}{I}$ e $\rho = \frac{\beta}{\alpha+\beta}$, obtemos $r = \frac{1/I}{p/\overline{b}+1/I} \times p = \frac{\overline{b}}{pI+\overline{b}} \times p$.

Para a obtenção de fluxos caracterizados por um envelope (b, r, p), submetemos cada fonte de tráfego a uma série de dois *leaky buckets* (figura 6). A geração de *tokens* é um processo determinístico com a chegada de um novo *token* a cada $1/r$ unidades de tempo no primeiro *bucket* e a cada $1/p$ no segundo. Na verdade, a taxa de geração de *tokens* no primeiro estágio determina a taxa média com a qual os pacotes entram na rede, enquanto o do segundo estágio limita a taxa de pico a um valor máximo de p. Devido as perdas no primeiro *leaky bucket*, a taxa média obtida r_{OUT} é menor do que r.

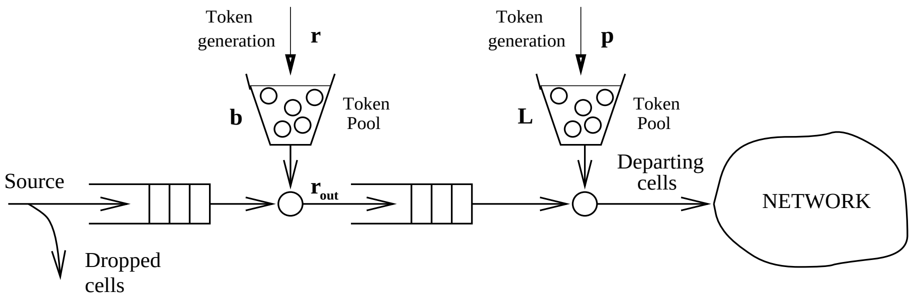


Figura 6: Dois *leaky buckets* em série.

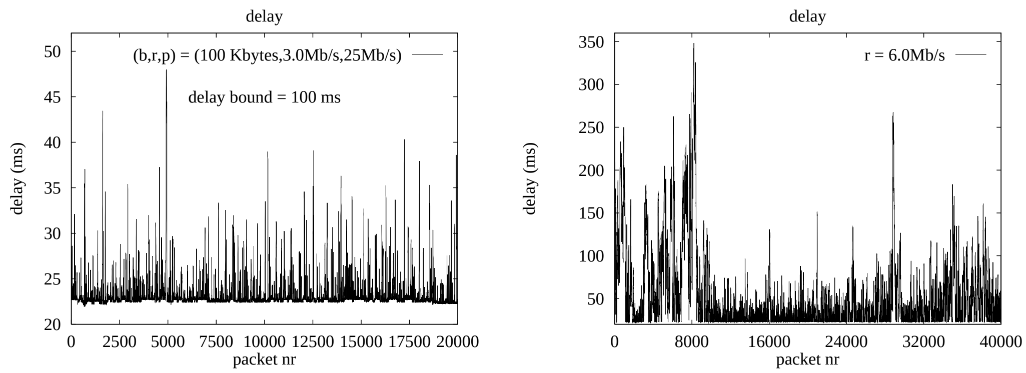
Cada fluxo GS apresenta a seguinte caracterização: $b = 100Kbytes$, $r = 3.0Mb/s$ e $p = 25Mb/s$. Neste cenário, as perdas observadas no primeiro *leaky bucket* giram em torno de 6%, então a verdadeira taxa média (r_{OUT}) está em torno de 2.82 Mb/s. O requisito de limite no retardo de transferência para os fluxos GS é de 100 ms. As taxas de limpeza R e os tamanhos do buffer necessários para satisfazer a esta exigência são dados na tabela 1. Os fluxos BR e BE cruzam todos os cinco comutadores. Todos os enlaces inter-comutadores e de acesso têm a banda passante de 51 Mb/s e são considerados como livres de erro.
Para verificar a viabilidade de garantir uma taxa média aos fluxos BR, usamos 5 fluxos

tamanho	R (Mb/s)	peso	B (# células)
5	8.4603	0.16589	1804
4	8.4562	0.16581	1804
3	8.4521	0.16573	1804
2	8.4480	0.16565	1804

enlace inter-comutador	R total (Mb/s)	porção reservada do enlace (%)
S_1-S_2	50.7372	99.49
S_2-S_3	50.7454	99.50
S_3-S_4	50.7454	99.50
S_4-S_5	50.7372	99.49

Tabela 1: taxa de limpeza (R), peso (W), tamanho do buffer (B), utilização do enlace.

BR com taxas média de 12Mb/s, 9Mb/s e 6Mb/s para os três primeiros, e 1.5 Mb/s para os dois últimos. Desta maneira, elevamos a utilização do enlace próxima do máximo, mas ainda respeitando a desigualdade expressa na equação 3. Assim, o peso associado a cada fluxo BR é $(\alpha_i \times 4.66e - 3)$ $, i = 1..5$, onde α_i representa o compartilhamento ponderado da banda passante deixada livre pelos fluxos GS depois de terem as suas respectivas taxas de limpeza reservadas.



(a) retardo do fluxo GS

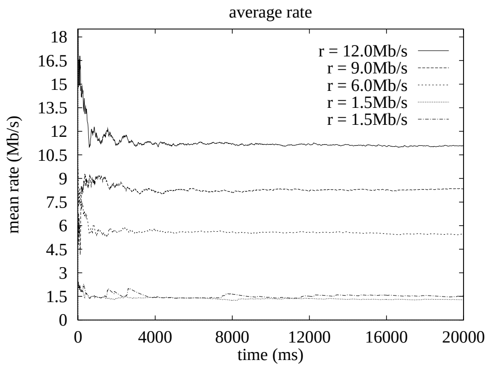
(b) retardo do fluxo BR

Figura 7: retardo de transferência.

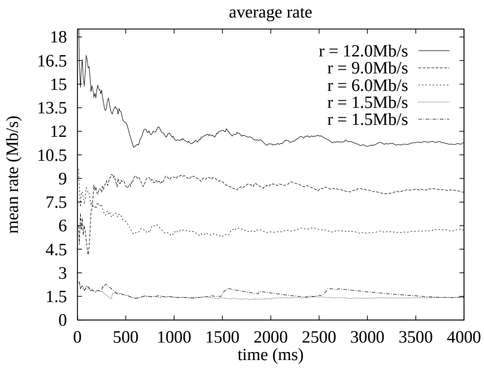
Os resultados de simulação são mostrados nas figuras 7 e 8. A curva 7(a) mostra o retardo experimentado pelos pacotes GS pertencentes a um único fluxo. Esta curva demonstra que os fluxos GS não são afetados pelos fluxos BR, uma vez que o limite no retardo de transferência de 100 ms é sempre respeitado. A curva 7(b) ilustra o comportamento do retardo de transferência de um fluxo BR. Como esperado, os pacotes do fluxo BR sofrem um retardo muito maior do que aquele experimentado por pacotes GS. As curvas 8(a) e 8(b) mostram, respectivamente, as taxas médias a curto e longo prazo alcançadas pelos cinco fluxos BR. A primeira ilustra como a taxa média dos fluxos BR é garantida a longo prazo. A segunda mostra que as taxas médias convergem rapidamente (depois de 3000 ms) aos valores desejados.

5 Conclusões

A contribuição deste artigo é dupla. Primeiramente, nós propomos uma arquitetura de integração de serviços para a comutação IP. Esta arquitetura objetiva principalmente fornecer



(a) taxas médias a longo prazo dos fluxos BR



(b) taxas médias a curto prazo dos fluxos BR

Figura 8: taxa média.

uma integração flexível e efetiva da QoS de cada fluxo com a rapidez dos equipamentos de comutação IP. A finalidade é construir uma estrutura capaz de responder aos requisitos cada vez mais restringentes das novas aplicações sobre a Internet. Num segundo tempo, definimos uma estratégia baseada na política de escalonamento WFQ para o suporte de uma nova classe de serviço (BR). O objetivo primordial desta classe é recuperar a porção não utilizada da banda passante reservada aos fluxos GS, sem afetar os retardos máximos de transferência dos pacotes destes fluxos. Os fluxos desta nova classe são menos sensíveis ao retardo de transferência, tais como o tráfego Web, e recebem garantias a longo prazo quanto à taxa média. Tendo em vista os resultados satisfatórios obtidos via simulação, o próximo passo no nosso trabalho de pesquisa consiste na implementação de um protótipo do nosso modelo teórico.

Referências

[1] S. Shenker and J. Wroclawski, “Network Element Service Specification Template,” RFC 2216, Sept. 1997.

[2] S. Shenker and J. Wroclawski, “General Characterization Parameters for Integrated Service Network Elements,” RFC 2215, Sept. 1997.

[3] J. Wroclawski, “The Use of RSVP with IETF Integrated Services,” RFC 2210, Sept. 1997.

[4] A. F. T. Committee, “Traffic Management Specification, version 4.0,” tech. rep., ATM Forum, Apr. 1996.

[5] P. Newman, G. Minshall, and T. Lyon, “IP Switching: ATM under IP,” *submitted to IEEE/ACM Transactions on Networking.*

[6] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, “RSVP: A New Resource ReSerVation Protocol,” *IEEE Network Magazine*, pp. 8–18, Sept. 1993.

[7] P. Newman, W. L. Edwards, R. Hinden, E. Hoffman, F. C. Liaw, T. Lyon, and G. Minshall, “Ipsilon’s General Switch Management Protocol Specification Version 1.1,” RFC 1987, Aug. 1996.

[8] P. Newman, W. L. Edwards, R. Hinden, E. Hoffman, F. C. Liaw, T. Lyon, and G. Minshall, “Ipsilon’s Flow Management Protocol Specification for IPv4 Version 1.0,” RFC 1953, May 1996.

[9] S. Shenker, C. Partridge, and R. Guérin, “Specification of Guaranteed Quality of Service,” draft-ietf-intserv-guaranteed-svc-08, Internet Draft, Feb. 1997.

[10] L. Georgiadis, R. Guérin, V. Peris, and R. Rajan, “Efficient Support of Delay and Rate Guarantees in an Internet,” in *Proceedings of ACM SIGCOMM’96*, (Stanford, CA), Aug. 1996.

[11] S. Lin and N. McKeown, “A Simulation Study of IP Switching,” in *Proceedings of ACM SIGCOMM’97*, (Cannes, France), Sept. 1997.