Estabelecimento de Reserva de Recursos para Fluxos *Multicast* em Túneis IP-sobre-IP *

Paulo Cesar S. Vidal[1] **, Otto Carlos M. B. Duarte**[2]

[1]Departamento de Engenharia de Sistemas – Instituto Militar de Engenharia
Praça General Tibúrcio, 80 – CEP 22290-270 - Rio de Janeiro, RJ

[2]Grupo de Teleinformática e Automação – Universidade Federal do Rio de Janeiro
COPPE/EE - Programa de Engenharia Elétrica
Caixa Postal 68504 – CEP 21945-970 - Rio de Janeiro, RJ

vidal@de9.ime.eb.br, otto@gta.ufrj.br

Resumo. *RSVP Túnel Multicast é um mecanismo de agregação de estados do protocolo RSVP utilizando túneis IP-sobre-IP. O mecanismo apresenta uma longa latência no estabelecimento de uma reserva de recursos, por causa da perda de mensagens RSVP, prejudicando o desempenho de aplicações multimídia. Este artigo aborda o emprego de temporizadores de renovação em etapas no mecanismo RSVP Túnel. Foram realizadas simulações para avaliar o mecanismo na presença de perdas de mensagens e os resultados mostram uma significativa diminuição no tempo de estabelecimento de sessões multicast com o uso de temporizadores de renovação em etapas.*

Abstract. *Multicast Tunnel RSVP is a state aggregation mechanism of the RSVP protocol using IP in IP tunneling. The mechanism presents a very long latency at reservation establishment because of the loss of RSVP messages. The problem reduces the performance of multimedia applications. This paper reports the use of staged refresh timers in the Tunnel RSVP mechanism at the set-up procedure of the resource reservation for multicast flows. The simulation results show a significant decrease of the establishment time of multicast sessions with the use of the staged refresh timers.*

1. Introdução

Muitas aplicações multimídias de tempo real estão sendo desenvolvidas para *Internet*, onde o modelo de encaminhamento de melhor esforço do protocolo IP (*Internet Protocol*) é inadequado. Portanto, existe a necessidade de fornecer às aplicações multimídias, classes de serviços com melhor qualidade (QoS - *Quality of Service*) em relação à largura de banda, atraso nas filas e perdas de pacotes. Buscando atingir estes objetivos, o IETF (*Internet Engineering Task Force*) definiu uma arquitetura de serviços integrados da *Internet* [Braden et al., 1994], onde são definidas classes de serviços que, se suportado pelos roteadores, podem fornecer um fluxo de dados com garantias de QoS. O nível de QoS fornecido por estas classes é programável por fluxo, de acordo com os requisitos das aplicações. Estes requisitos podem ser passados para os roteadores, usando o protocolo RSVP (*Resource Reservation Protocol*) [Braden et al., 1997].

Um fluxo é identificado pelo endereço *unicast* ou *multicast* e seu estado é armazenado nos roteadores entre a fonte e o destino do fluxo de dados. A larga utilização dos serviços integrados, significa uma grande quantidade de estados armazenados nos roteadores e um grande

*This work was supported by IME, UFRJ, FUJB, PROTEM-CC, CNPq and CAPES.

processamento para gerenciá-los, podendo sobrecarregar os roteadores. Consequentemente, a técnica de agregação de fluxos tem sido utilizada para alcançar a escalabilidade necessária para fornecer os serviços integrados, reduzindo a quantidade de estados armazenados para fluxos *unicast* ou para fluxos *multicast*. Estudos tem sido realizados para alcançar a redução de estados e mecanismos de agregação tem sido propostos.

Em [Terzis et al., 1998], o protocolo RSVP é utilizado em túneis IP-sobre-IP onde várias sessões individuais ponto-a-ponto podem ser mapeadas para uma sessão túnel entre os dois pontos da extremidade do túnel. O mecanismo de agregação é denominado de RSVP Túnel *unicast*. Os roteadores interiores somente armazenam estados das sessões RSVP túneis. Em [Baker et al., 2001] descreve o uso de uma única reserva RSVP para agregar outras reservas em uma região de agregação, de modo conceitualmente similar ao uso de caminhos virtuais em uma rede ATM. O mecanismo cria dinamicamente uma reserva agregada, classifica o tráfego para qual o agregado se aplica, e determina a largura de banda necessária para o agregado.

O trabalho de [Vidal and Duarte, 2000] apresenta um mecanismo de agregação de estados do protocolo RSVP para fluxos *multicast* utilizando a técnica de tunelamento IP-sobre-IP. O mecanismo é uma extensão do RSVP Túnel *unicast*, realizando o mapeamento de uma sessão *multicast* em *n* sessões túneis. O mecanismo apresenta uma longa latência no estabelecimento de uma reserva de recursos, por causa da perda de mensagens RSVP. Se a primeira mensagem RESV Túnel enviada pelo roteador for perdida na região de agregação, uma cópia da mensagem não será retransmitida até o término do intervalo de renovação, ocasionando um atraso de 30 s ou mais até que a reserva seja estabelecida. Este problema não deve ser aceitável para determinadas aplicações multimídia.

Este artigo aborda o emprego de temporizadores de renovação em etapas (*staged refresh timers*) no mecanismo RSVP Túnel, para configuração de reserva de recursos de fluxos *multicast*. Temporizadores em etapas controlam o intervalo de retransmissão das mensagens RSVP túneis. O temporizador inicia com um pequeno intervalo de tempo de retransmissão no qual vai aumentando exponencialmente até alcançar um limite. A diminuição no tempo de retransmissão de mensagens RSVP ocasiona tempos menores na configuração de reservas de recursos. Foram realizadas simulações para avaliar o mecanismo na presença de perdas de mensagens e os resultados mostram uma significativa diminuição no tempo de estabelecimento de sessões *multicast* com o uso de *staged refresh timers*.

O artigo está organizado da seguinte forma. A Seção 2 apresenta o processo de estabelecimento de reservas de recursos do protocolo RSVP e a Seção 3 descreve o modelo de túneis RSVP para fluxos *unicast*. A Seção 4 descreve o mecanismo RSVP Túnel *Multicast* para agregar fluxos *multicast*. A Seção 5 apresenta resultados da simulação na utilização do *staged refresh timers* no mecanismo de agregação RSVP Túnel *Multicast*. A Seção 6 mostra considerações finais e trabalhos futuros.

2. Estabelecimento de Reservas de Recursos no protocolo RSVP

O protocolo RSVP pode ser utilizado por uma estação para solicitar reserva de recursos, através da classe de serviço garantido ou de carga controlada, para fluxos de dados de uma aplicação *unicast* ou *multicast*. As principais mensagens usadas pelo RSVP são as mensagens *Path*, transmitidas pelos emissores, e as mensagens *Resv*, transmitidas pelos receptores. A mensagem *Path* é usada para armazenar o estado do caminho em cada roteador entre a fonte e o destino do fluxo. O estado do caminho é usado para encaminhar mensagens *Resv* na direção inversa. A mensagem *Resv* conduz o pedido de reserva para os roteadores ao longo do caminho entre o emissor e o receptor.

O estado de uma determinada reserva é mantido pelos roteadores através do mecanismo *soft-state*. Para cada estado da reserva está associado um temporizador, que controla o tempo de duração do estado. O protocolo RSVP mantém atualizado o estado através do envio periódico de mensagens *Path* e mensagens *Resv* entre a fonte e o destino em intervalos de 30 s. Se após um certo período de tempo (*time-out*) um roteador não receber, por exemplo, uma mensagem *Resv* de determinado destino, ele libera os recursos que haviam sido alocados para o receptor que enviou a mensagem.

O mecanismo *soft-state* não fornece confiabilidade na entrega de mensagens e pode causar problemas na ocorrência de perda de mensagens. Perdas de pacotes na *Internet* podem ser frequentes em alguns enlaces. Medidas realizadas no MBone (*Internet Multicast Backbone)* [Yajnik et al., 1996] apresentaram uma taxa média de perda de pacotes de aproximadamente 1 a 2% e ocasionalmente alcançar 20% em enlaces com congestionamento. Se a primeira mensagem *Path* ou *Resv* de um usuário for perdida na rede, uma cópia da mensagem não será retransmitida até o término do intervalo de renovação, ocasionando um atraso de 30 s ou mais no estabelecimento de uma reserva.

Para reduzir o tempo e tornar confiável o processo de estabelecimento de reservas na presença de perdas de mensagens RSVP, foi proposto o uso de temporizadores de renovação em etapas (*staged refresh timers*) [Pan and Schulzrinne, 1997]. Após enviar uma mensagem *Path* inicial, um nó RSVP espera um período de tempo igual ao período inicial de estabelecimento T_0. Se após este tempo uma mensagem de reconhecimento *PathAck* não foi recebida, o nó retransmite a mensagem *Path*. Para ser adaptativo em diversas situações de congestionamento, o valor do período de estabelecimento deve ser multiplicado pelo fator $(1 + \Delta)$ a cada retransmissão de uma mensagem *Path*. A retransmissão em etapas continuará até que uma mensagem *PathAck* seja recebida ou o intervalo de retransmissão alcance um valor limite R_l. O mesmo processo se aplica na transmissão de mensagens *Resv*. Os valores usados para T_0, Δ e R_l são respectivamente 3s, 0.3 e 30s.

3. Agregação em Túneis IP-sobre-IP

A técnica de tunelamento IP-sobre-IP pode ser usada para criar uma região de agregação de fluxos. Desta forma, os fluxos de múltiplas fontes que chegam ao roteador de entrada da região de agregação são encaminhados, de uma forma agregada, para o roteador de saída da região de agregação em um túnel IP-sobre-IP. Assim, todos os roteadores internos à região de agregação processam um fluxo agregado ao invés de processar os fluxos individualmente.

Baseado neste princípio, o trabalho de [Terzis et al., 1998] implementa um mecanismo de agregação de estados do protocolo RSVP sobre túneis para fluxos *unicast*.

O mecanismo de túneis RSVP está representado na Figura 1, onde a fonte e o receptor de uma sessão RSVP individual (fim-a-fim) estão conectados pelo túnel. O roteador $R_{entrada}$ é um roteador de entrada do túnel que encapsula e encaminha os datagramas para o roteador R_{saida} através dos roteadores intermediários. O roteador R_{saida} é o ponto final do túnel que recebe, desencapsula e encaminha os datagramas baseado no cabeçalho original.

No mecanismo existem duas sessões diferentes, uma sessão RSVP túnel e uma sessão RSVP fim-a-fim. A sessão RSVP fim-a-fim estabelece uma ligação com reserva de recursos entre a fonte e o receptor. A sessão RSVP túnel localizada entre o ponto $R_{entrada}$ e o ponto R_{saida} fornece reserva de recursos para um conjunto de sessões RSVP individuais configu-

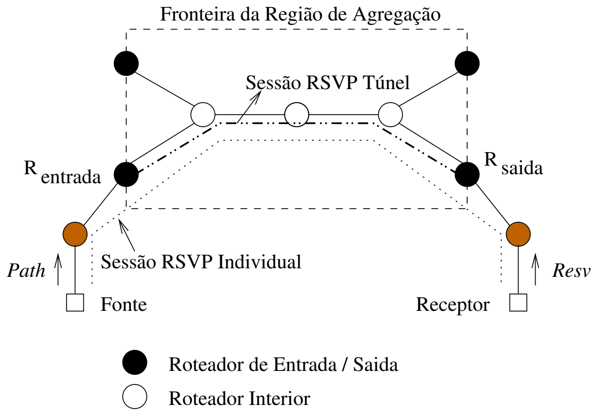

Figura 1: Modelo de túneis RSVP *unicast*.

radas no túnel.

As mensagens *Path* e *Resv* de uma sessão fim-a-fim que chegam nos pontos finais do túnel são mapeadas para uma sessão túnel correspondente à uma classe de tráfego agregada. Elas são encapsuladas e desencapsuladas na mesma maneira que pacotes IP, sendo acrescidas do cabeçalho externo especificando o ponto de entrada do túnel como fonte e o ponto de saída como destino.

A configuração de reservas de recursos no túnel é realizada através dos seguintes passos:

1. o roteador $R_{entrada}$ recebe uma mensagem *Path* individual e cria o estado do caminho (*Path*) da sessão RSVP fim-a-fim;

2. o roteador $R_{entrada}$ envia a mensagem *Path* fim-a-fim encapsulada sobre o túnel para o roteador R_{saida};

3. o roteador R_{saida} recebe a mensagem, desencapsula, cria o estado do caminho da sessão RSVP fim-a-fim e encaminha a mensagem original ao destino;

4. quando uma mensagem *Resv* fim-a-fim correspondente à sessão RSVP fim-a-fim chega no roteador R_{saida}, o roteador cria o estado de reserva, encapsula e envia a mensagem para o roteador de entrada;

5. o roteador de entrada recebe a mensagem *Resv*, desencapsula, cria o estado de reserva e mapeia a sessão RSVP fim-a-fim para uma sessão RSVP túnel;

6. após o mapeamento, o roteador de entrada envia uma mensagem PATH túnel correspondente à sessão RSVP túnel;

7. o roteador R_{saida} recebe a mensagem PATH túnel e envia uma mensagem RESV túnel para o interior da região de agregação, solicitando reserva de recursos;

8. depois de receber a mensagem RESV túnel, o roteador de entrada envia uma mensagem RESV_CONFIRM túnel para o roteador R_{saida} como confirmação da reserva no túnel;

9. o roteador de saída recebe a mensagem RESV_CONFIRM túnel e envia a mensagem *Resv* encapsulada para o roteador de entrada;

10. o roteador de entrada recebe a mensagem *Resv* encapsulada, e encaminha a mensagem *Resv* original para a fonte.

4. Mecanismo de Agregação *Multicast*

O mecanismo de agregação de estado dos fluxos *multicast* [Vidal and Duarte, 2000] é aplicado ao serviço de carga controlada e agrega o tráfego em trânsito na região de agregação

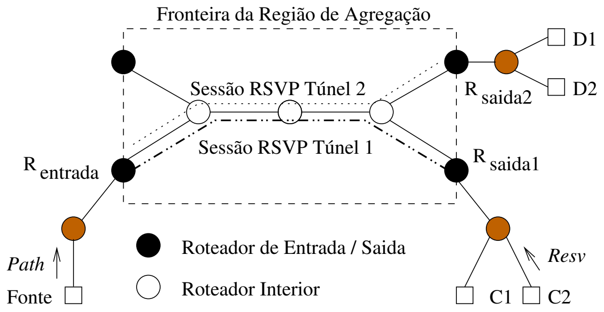

Figura 2: Mecanismo de túneis RSVP *multicast*.

usando a técnica de tunelamento.

O mecanismo é uma extensão do modelo de túneis [Terzis et al., 1998] e as principais mudanças introduzidas foram o mapeamento de uma sessão RSVP individual para várias sessões RSVP túnel, a difusão de uma mensagem *Path* para vários roteadores de saída e a marcação dos pacotes reservados através do campo TOS (*Type of Service*) ou o campo DS (*Differentiated Services*) do cabeçalho externo do pacote IP.

A associação de uma sessão RSVP *multicast* para várias sessões RSVP túneis está ilustrado na Figura 2. A Figura mostra uma árvore *multicast* composta de uma fonte e quatro receptores (C1,C2 no domínio C e D1, D2 no domínio D), todos pertencentes à sessão *multicast*. No exemplo foram criadas duas sessões RSVP túnel localizadas entre o roteador $R_{entrada}$ e os roteadores R_{saida1} e R_{saida2} para fornecer a reserva de recursos sobre cada túnel à sessão *multicast* RSVP fim-a-fim.

O roteador de entrada inicia o processo de configuração da reserva em cada túnel através das seguintes etapas:

1. recebe uma mensagem *Path* individual e cria o estado do caminho (*Path*) da sessão RSVP fim-a-fim;

2. envia uma mensagem *Path* fim-a-fim encapsulada para o roteador R_{saida1} e para o roteador R_{saida2};

3. para cada sessão RSVP túnel, são executados passos 3 a 10 descritos na Seção 3. da configuração da reserva de recursos de uma sessão RSVP túnel *unicast*.

Os temporizadores de renovação em etapas proposto em [Pan and Schulzrinne, 1997] fornecem confiabilidade através de mensagens de reconhecimento *PathAck* e *ResvAck*. Entretanto, no mecanismo RSVP Túnel uma mensagem RESV túnel representa um reconhecimento de uma mensagem PATH túnel e uma mensagem RESV_CONFIRM túnel representa um reconhecimento de uma mensagem RESV túnel. Portanto não necessita incluir novas mensagens no mecanismo para ter confiabilidade. Quando o roteador R_{saida} envia uma mensagem RESV túnel, ele inclui um objeto RESV_CONFIRM túnel na mensagem e em uma lista de espera de mensagens RESV_CONFIRM túnel. O objeto contém o endereço do roteador de entrada e o número do objeto. Existe um temporizador de renovação em etapas associado a cada objeto. Caso não receba a mensagem RESV_CONFIRM túnel, uma mensagem RESV túnel é retransmitida usando o processo descrito na Seção 2.

Os pacotes de dados reservados são encapsulados IP-sobre-IP e pelo protocolo UDP (*User Datagram Protocol*) para diferenciá-los de pacotes não reservados [Terzis et al., 1998]. Entretanto o uso do encapsulamento UDP aumenta o tempo de processamento nos roteadores. Assim, os pacotes de dados referentes à sessão *multicast* que chegam ao ponto $R_{entrada}$

e que possuem reservas de recursos associada às duas sessões túnel podem ser marcados com a classe de tráfego no campo TOS do cabeçalho externo do pacote IP, ou pelo campo DS [Nichols et al., 1998]. A utilização do campo DS é para manter a compatibilidade com pedidos de reserva que usam Serviços Diferenciados. Após a marcação, são geradas cópias (uma para cada túnel), sendo cada cada pacote é encapsulado e enviado para o endereço do roteador de saída. Quando chegam nos roteadores borda de saída são desencapsulados e enviadas para o destino.

O número de cópias do pacote inseridas na região aumenta linearmente com o número de roteadores de saída, assim aumentando a largura de banda consumida pela rede. Uma solução para o problema é usar um endereço *multicast* reservado no campo endereço de destino do cabeçalho externo do pacote IP, em vez de usar o endereço do roteador de saída. O endereço *multicast* representaria um conjunto de roteadores de saída. Cada roteador na região deve ter uma tabela com dois campos: endereço *multicast* reservado e a lista de roteadores de saída. O estado de cada sessão túnel seria mantido. O roteador de entrada envia um pacote e sua duplicação ocorre no roteador que tem duas interfaces.

5. Resultados

Nesta seção são apresentados os resultados obtidos por simulações do mecanismo de agregação *multicast*. O objetivo da simulação é comparar a funcionalidade do mecanismo na presença de perda de mensagens RSVP Túnel utilizando temporizadores de renovação de taxa fixa e temporizadores de renovação em etapas. Para obtenção das medidas foi utilizado o simulador de redes de computadores *ns* (*Network Simulator*). Uma extensão ao pacote básico do *ns* foi implementada para fornecer o suporte ao mecanismo de agregação sendo constituída de dois módulos: RSVP clássico (fluxos individuais) e RSVP Túnel (fluxos agregados). O módulo RSVP clássico foi obtido de [Greis, 1999].

O módulo RSVP Túnel desenvolvido executa as seguintes funções: processamento das mensagens PATH , RESV e RESV_CONFIRM Túnel; mapeamento das sessões individuais para sessões túnel; encapsulamento e desencapulamento dos pacotes; manutenção das tabelas de estados agregados e retransmissão de mensagens usando *staged refresh timers*.

As simulações consistiram de repetidos estabelecimentos de reservas de recursos em grupos *multicast* (um grupo é constituído de uma fonte e *n* receptores), sobre vários tamanhos de árvores *multicast* e várias condições de perda de pacotes.

O modelo de perda é baseado na taxa de erro por pacotes, isto é, no percentual de pacotes perdidos em cada direção do enlace. Existe dois tipos de enlace quanto a taxa de erro: enlace com baixa taxa de erro (0.1 a 1% de perdas) e enlace com média taxa de erro de pacotes (2 a 5% de perdas).

Foram utilizadas dois tipos de topologias nas simulações: uma árvore binária cheia e uma árvore binária com enlace de gargalo (*bottleneck link*). As configurações da árvore binária cheia são com 3, 7 e 15 nós e as configurações da árvore binária com gargalo são com 4, 8 e 16 nós. As Figuras 3(a) e 3(b) apresentam uma árvore binária cheia com 7 nós e uma árvore binária com enlace de gargalo com 8 nós. O nó raiz é o roteador de entrada e os nós folha são os roteadores de saída.

Para cada configuração, foram estabelecidos 2000 fluxos com reservas de recursos. Não foi introduzido atraso nos enlaces para isolar o tempo de processamento da operação do mecanismo. O tempo médio de duração de cada sessão *multicast* é de 180s com distribuição exponencial. A tempo médio de transmissão entre as mensagens PATH túnel é de 100ms

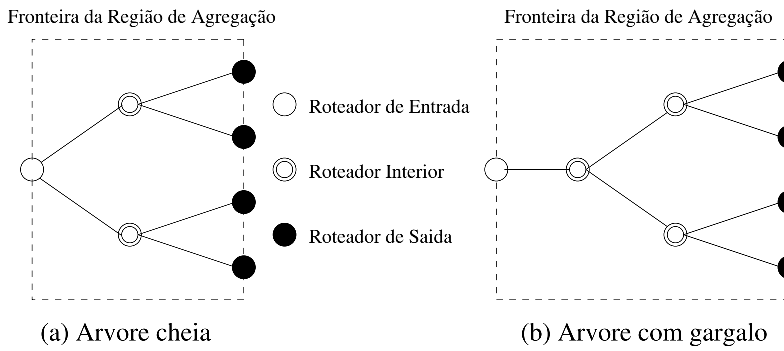

Figura 3: Topologias utilizadas nas simulações.

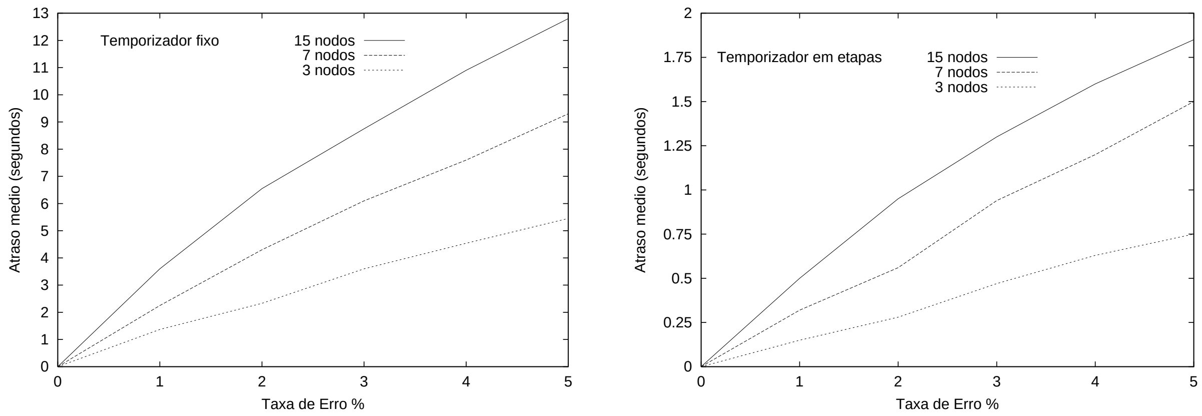

Figura 4: Atraso médio na ArvBin cheia com 1-5% de perda de pacotes.

com distribuição exponencial. Uma mensagem RESV túnel é enviada imediatamente após a chegada da mensagem PATH túnel. O valor da taxa fixa de renovação de mensagens é de 30s. Os valores usados para T_0, Δ e de R_l são respectivamente 3s, 0.3 e 30s.

As Figuras 4 e 5 mostram o atraso médio de estabelecimento de sessões com diversas taxas de erro por enlace e o ganho em desempenho usando os temporizadores em etapas tanto na árvore binária cheia quanto na árvore binária com enlace de gargalo. O eixo horizontal do gráfico representa o percentual de pacotes perdidos (taxa de erro) por enlace. Cada atraso médio é resultante de um conjunto de simulações realizadas em uma topologia, onde uma taxa de erro é inserida em todos os enlaces da topologia.

Note um atraso muito pequeno no estabelecimento de sessões com temporizadores em etapas em enlaces com taxa de 3-5% de perda na árvore de 3 e 4 nós (Figuras 4 e 5). Este resultado é de particular interesse no caso de uma mudança de rota, pois o roteador na atualização da rota pode produzir rajadas de mensagens de controle na qual temporariamente pode congestionar o enlace. Na topologia com enlace de gargalo (taxa de erro 2-5%) houve um pequeno aumento (menor que 1s) no atraso médio em relação a topologia sem gargalo, o que mostra a eficiência do mecanismo em enlaces de gargalo com média taxa de erro.

6. Considerações Finais

O emprego de *staged refresh timers* no mecanismo RSVP túnel *multicast* apresentado

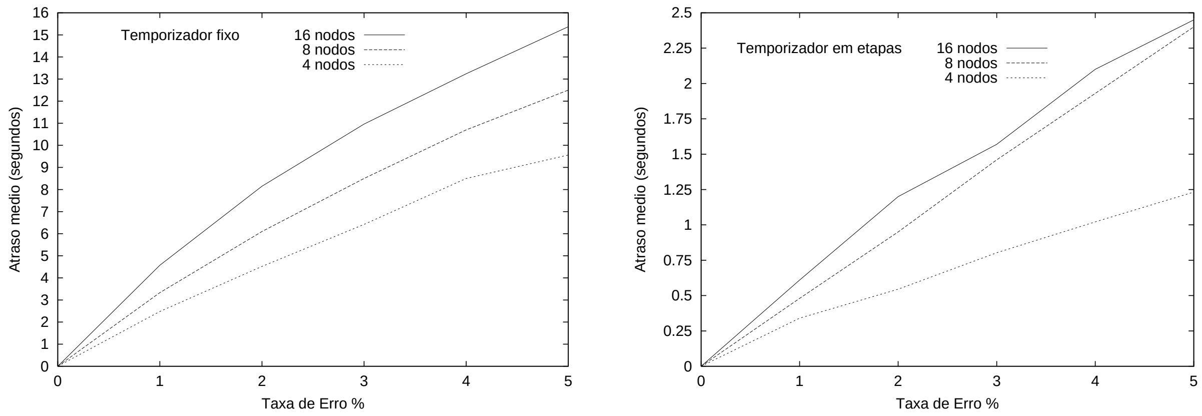

Figura 5: Atraso médio na ArvBin com gargalo com 1-5% de perda.

neste artigo mostrou à redução no tempo de estabelecimento de reservas na presença de perdas de mensagens RSVP. O mecanismo RSVP Túnel apresenta confiabilidade quando uma mensagem RESV túnel representa um reconhecimento de uma mensagem PATH túnel e uma mensagem RESV_CONFIRM túnel representa um reconhecimento de uma mensagem RESV túnel. Portanto não precisa incluir novas mensagens no mecanismo para ter confiabilidade. Como trabalhos futuros podemos citar a realização de simulações em topologias maiores com um número maior de fluxos sob diferentes diferentes condições de congestionamento.

Referências

Baker, F., Iturralde, C., Faucheur, F. L., and Davie, B. (abril de 2001). Aggregation of RSVP for IPv4 and IPv6 Reservations. *Internet-Draft*.

Braden, R., Clark, D., and Shenker, S. (julho de 1994). Integrated Services in the Internet Architecture: An Overview. *RFC 1633*.

Braden, R., Zhang, L., Berson, S., and Jamin, S. (setembro de 1997). Resource Reservation Protocol (RSVP) - Version 1 functional specification. *RFC 2205*.

Greis, M. (1999). RSVP/ns: An Implementation of RSVP for the Network Simulator ns-2. *http://titan.cs.uni-bonn.de/~greis/rsvpns/rsvpns.ps.gz*.

Nichols, K., Blake, S., Baker, F., and Black, D. L. (dezembro de 1998). Definition of the differentiated services field (DS field) in the IPv4 and IPv6 headers. *RFC 2474*.

Pan, P. and Schulzrinne, H. (novembro de 1997). Staged Refresh Timers for RSVP. *Proceedings of Global Internet*.

Terzis, A., Zhang, L., and Hahne, E. L. (maio de 1998). Reservations for Aggregate Traffic: Experiences from an RSVP Tunnels Implementation. *6th International Workshop on Quality of Service (IWQoS'98)*.

Vidal, P. C. S. and Duarte, O. C. M. B. (setembro de 2000). Agregação de Estados do Protocolo RSVP para Fluxos *Multicast*. *XVIII Simpósio Brasileiro de Telecomunicações*.

Yajnik, M., Kurose, J., and Towsley, D. (novembro de 1996). Packet Loss Correlation in the MBone Multicast Network. *Proceedings of Global Internet*.