
Qualidade de Serviço

Pedroso

4 de março de 2009

1 Introdução

A implantação de qualidade de serviço (QoS, Quality of Service) na rede é essencial para o funcionamento apropriado de diversas aplicações, como videoconferência e VoIP (voz sobre IP). Estas aplicações demandam, além de grande largura de banda, um serviço diferenciado. Em muitos casos, é preciso garantir que a transmissão de dados seja feita sem interrupção ou perda de pacotes.

O Protocolo IP não Privilegiava o QoS Ao ser criado, o protocolo IP tinha como meta a comunicação com controle de tráfego baseado na regra do menor esforço, que não provê mecanismos de qualidade de serviço (QoS), o que compromete a alocação de recursos da rede. Naquela época não se previa o crescimento explosivo da Internet, muito menos a integração de voz e dados numa única rede de pacotes: a rede IP. Isso fez surgir a necessidade do desenvolvimento de protocolos que garantissem a qualidade de serviço fim a fim, uma vez que voz e vídeo são muito sensíveis à qualidade da transmissão. A qualidade de serviço é definida pelo IETF (Internet Engineering Task Force, <http://www.ietf.org>) como o conjunto de requerimentos que devem ser satisfeitos quando do transporte de um fluxo. Desta forma, a qualidade de serviço pode ser vista como uma garantia de desempenho em relação a determinadas métricas, principalmente:

1. Taxa de transmissão;
2. Atraso fim a fim (ver Figura 1);
3. Jitter;
4. Perda de pacotes;

Ver exemplo em sala de aula sobre a importância de cada uma destas métricas em uma transmissão VoIP.

Leitura adicional:

<http://www.rnp.br/newsgen/9911/qos.html>
http://www.rnp.br/newsgen/0005/qos_voip1.html
<http://www.rnp.br/newsgen/0005/rsvp.html>

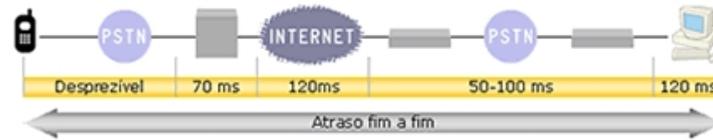


Figura 1: Atraso Fim a Fim

2 Algoritmos de Escalonamento de Pacotes

Pedidos de qualidade de serviço podem ser enviados através da rede utilizando-se o protocolo RSVP. O sistema operacional do roteador deve estar preparado para separar o tráfego e realizar o controle sobre as reservas efetuadas. Para realizar esta separação é necessário um algoritmo de escalonamento de pacotes que garanta que os fluxos tenham acesso a uma determinada parcela total do enlace, de acordo com as reservas estabelecidas. O algoritmo de escalonamento de pacotes utilizado originalmente no TCP/IP é o FIFOQ (First In First Out Queuing). Os mecanismos existentes originalmente para o controle de congestionamento são:

- Controle de fluxo baseado no mecanismo de janelas do TCP;
- Tratamento das mensagens ICMP e redução da taxa de transmissão espontaneamente pelo transmissor.

Os mecanismos acima não são adequados para garantir uma determinada qualidade de serviço e, como consequência, podemos observar casos de colapso total da rede durante períodos de alto tráfego, como estudado na RFC 1254. A RFC 1254 faz um levantamento de alguns métodos de controle de congestionamento existentes e divide o controle de congestionamento em duas categorias: algoritmos de prevenção e recuperação do congestionamento.

Deve-se enfatizar que o RSVP é responsável pelo estabelecimento e manutenção de um estado de reserva. O modo como esta reserva será implementada pelo roteador pode variar e depende do algoritmo de escalonamento de pacotes utilizado. Em 1995 foi proposto um modelo baseado em classes para separação dos diversos estilos de tráfego [36]. Este modelo é chamado CBQ (Class Based Queuing). Outros algoritmos disponíveis são o SFQ, WFQ, RED, entre outros.

2.1 First In First Out Queuing (FIFOQ)

Este é o algoritmo de escalonamento de pacotes utilizado originalmente no conjunto de protocolos Inter Rede. O algoritmo está baseado na criação de um buffer de tamanho fixo para cada interface física. Os pacotes que devem ser enviados por esta interface são armazenados neste buffer e enviados em ordem de chegada - o primeiro pacote a ser transmitido será o pacote mais antigo do buffer. Quando não há mais espaço no buffer, os pacotes que chegam são descartados. O problema deste método é a impossibilidade da separação do tráfego e divisão da banda passante entre os diversos fluxos IP.

2.2 GPS - Generic Processor Sharing

È o melhor escalonador possível. Seleciona n bytes de cada fila, onde n é o peso.

No entanto, não é implementável, pois os pacotes não podem ser divididos. Uma rede ATM usa este método utilizando como menor unidade a célula de tamanho fixo (53 bytes). Isto é possível porque no ATM todas as células possuem o mesmo tamanho.

2.3 WFQ - Weighted Fair Queue

Os pesos são colocados do mesmo modo que no WRR. No entanto, o WFQ simula o algoritmo GPS bit a bit para descobrir qual pacote terminaria antes sua transmissão. O pacote que terminar antes na simulação será transmitido. Prova-se que o WFQ é algoritmo de escalonamento mais justo para uma rede de pacotes de tamanho variável. O problema é o desempenho computacional pobre para a implantação prática.

2.4 WRR - Weighted Round Robin

Cada reserva irá criar uma fila no elemento de rede. O escalonador seleciona pacotes das filas em uma seqüência circular entre as filas. A quantidade de pacotes selecionados será igual ao peso da fila. Exemplo: 3 filas A, B, C, peso 1,3 e 2. O tamanho médio do pacote será 100, 200 e 1000 bits respectivamente pra as filas A, B e C.

A divisão de banda obtida será:

$$|x| = \begin{cases} A = \frac{100}{100+3.200+2.100}=0,018 \\ B = \frac{3.200}{100+3.200+2.100}=0,592 \\ C = \frac{2.100}{100+3.200+2.100}=0,388 \end{cases}$$

2.5 Class Based Queuing (CBQ)

O CBQ foi proposto por Jacobson e tem sido estudado por Floyd. Este método endereça o particionamento e divisão da banda disponível através de uma estrutura hierárquica de classes. Cada classe possui sua própria fila e sua banda. Uma classe filho pode emprestar da classe pai quando houver sobra de banda. Adicionalmente, o CBQ permite que seja utilizado um algoritmo de escalonamento de pacotes diferente gerenciando cada fila criada.

A Figura 2 mostra os componentes básicos do CBQ. O mecanismo de funcionamento é o seguinte: o classificador separa os pacotes chegando para sua própria classe. O estimador estima a banda recentemente utilizada pela classe. Se uma classe excedeu seu limite pré-definido, o estimador marca a classe como fora de limite. O escalonador determina o próximo pacote a ser enviado (entre todas as classes disponíveis), baseado nas prioridades e no estado das classes.

No CBQ, as classes são atendidas pelo algoritmo Round Robin. Floyd prova em que é vantajoso utilizar-se o WRR (Weighted Round Robin) ao invés do PRR (Packet by Packet Round Robin) para o serviço garantido.

No algoritmo Round Robin, cada classe poderá transmitir uma certa quantidade de dados. Após realizar a transmissão, o escalonador seleciona a classe seguinte e transmite uma certa quantidade

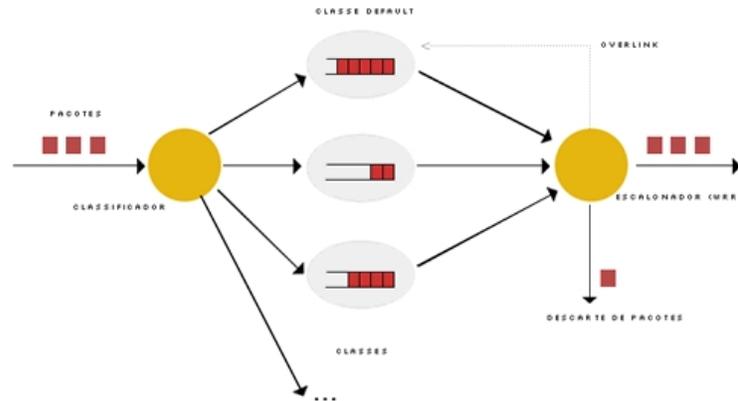


Figura 2: Componentes CBQ

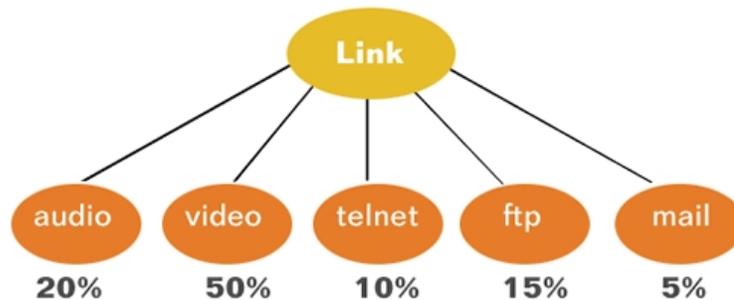


Figura 3: Alocação do Link

de dados desta classe. Quando uma classe está marcada como fora de limites, ela não será escalonada.

Em seu artigo, Floyd explica que protocolos diferentes podem ser tratados como classes diferentes, tendo alocação de banda diferenciada (ex. IP 40%, IPX 30%, SNA 30%). A visão de classe pode ser colocada para o protocolo IP, onde existem classes diferentes para cada serviço (por exemplo, como mostrado na Figura 3). Em implementações práticas deve existir uma classe default, que abriga o tráfego de serviços não especificados explicitamente na configuração do sistema.

2.6 Integração de Tráfegos

O grande desafio para integração de diferentes tráfegos na rede comutada de pacotes, como a Internet, reside em prover serviços adequados para tratar tanto o tráfego de dados como o de tempo real, como voz e vídeo. O modelo de serviço tradicional, baseado no menor esforço, é adequado para grande número de aplicações envolvendo transporte de dados, não tão sensíveis ao tempo. Entretanto, sensibilidade ao tempo (atrasos e variação de atrasos) é uma característica do transporte de voz e vídeo, e outros modelos de serviço tornam-se necessários para viabilizar esse tipo de tráfego.

Ou seja, para garantir qualidade de serviço para um determinado tipo de tráfego, deve-se conhecer antes, suas principais características. Assim sendo, o IETF apresentou duas propostas de modelos de serviços: o modelo de Serviços Integrados (IntServ) e o modelo de Serviços Diferenciados (Diff-Serv). Cada um desses modelos possui características apropriadas a situações distintas, sendo a combinação de ambos o caminho para a obtenção da qualidade de serviço procurada.

IntServ Os recursos da rede são divididos de acordo com as solicitações de QoS das aplicações e submetidos a um policiamento de tráfego no gerenciamento de banda.

DiffServ O tráfego é dividido em classes de serviço e os recursos da rede serão compartilhados dando um tratamento para cada uma dessas classes.

Características apropriadas Estes dois modelos de QoS podem ser aplicados em um único fluxo (per flow) fim-a-fim ou em um agregado de fluxos (per aggregate). Para a implementação do modelo IntServ o protocolo a ser utilizado é o ReSerVation Protocol (RSVP) que providência sinalização para a rede reservar recursos. Embora usado tipicamente em tráfego per flow, pode também ser usado em agregados.

Principais características Sabemos, por exemplo, que o tráfego de voz é sensível a atrasos e, em caso de tráfego real-time, seria melhor não entregar um segmento da informação (pacote) do que entregá-lo com atraso. O tráfego de dados, em sua grande maioria, tem a característica de não ser sensível a atrasos, mas pode ser sensível à perda de informações, que devem ser entregues, mesmo com atraso. Para garantir QoS faz-se necessário estabelecer um contrato entre a aplicação e a rede na qual fluirá o tráfego originado desta aplicação. A este contrato damos o nome de Contrato de Nível de Serviço, ou Service Level Agreement - SLA. O SLA é estabelecido sobre alguns parâmetros e, em caso da impossibilidade do cumprimento deste, a rede estará incapacitada de prover o referido serviço. Os parâmetros são: latência ou atraso, variação do atraso, largura de banda, perda de pacotes e disponibilidade.

3 Mecanismos de Qualidade de Serviço no IP

O protocolo IP é baseado em datagrama não confiável. Esta é uma das características mais importantes do conjunto de protocolos inter rede e foi adotado inicialmente devido a ser mais simples e robusto em comparação ao modelo de transmissão baseado em conexões. Como não são estabelecidas conexões e não existia garantia de atraso máximo de cada datagrama, o protocolo originalmente não suportava a transmissão de categorias de tráfego que exigem a garantia de serviço, ex. voz, vídeo - exceto em situações de baixo tráfego. A questão chave para a solução dos problemas é o estabelecimento de conexões capazes de garantir certos parâmetros de desempenho, estes parâmetros descrevem uma determinada qualidade de serviço (QoS - Quality of Service). Deve ser proposto um mecanismo capaz de estabelecer e manter reservas de recursos para uma determinada aplicação, o que introduz problemas não tratados no projeto original do IP. Consumo de Serviços Integrados A RFC 1633 especifica um modelo de serviços integrados para o conjunto de protocolos inter rede, propondo a criação de três categorias de serviço:

- Melhor Esforço: utilização da banda disponível;

- Serviço de Carga Controlada: garantia de uma determinada taxa média de transmissão, sem compromisso com relação ao atraso experimentado por cada pacote;
- Serviço Garantido: garantia de uma taxa mínima de transmissão e um atraso máximo determinístico experimentado por cada pacote.

Para tornar possível a implementação deste modelo, é introduzido o conceito de fluxo como sendo uma seqüência de pacotes envolvendo uma única comunicação simplex entre transmissor e receptor. O modelo definido pela RFC 1633 separa o comportamento esperado do método utilizado para a sua implementação. O modelo de implementação do controle de tráfego, que permite a separação dos fluxos em suas categorias de serviço, contém os seguintes módulos básicos:

- Escalonador de pacotes: estabelece qual será a política de escolha do próximo pacote a ser transmitido; detalhes sobre o escalonador de pacotes dependem de características do meio de transmissão e das funcionalidades fornecidas pelo protocolo de enlace utilizado;
- Classificador de pacotes: Verifica a qual categoria de tráfego pertence o pacote a ser transmitido, os fluxos podem ser divididos em classes de acordo com a similaridade do tratamento a ser dado ao fluxo. Por exemplo, todos os fluxos carregando voz podem ser agrupados na mesma classe de tráfego;
- Mecanismo de controle de admissão: Determina se uma nova reserva pode ou não ser estabelecida. Não se deve confundir o controle de admissão com o policiamento do tráfego, que é função do escalonador de pacotes. O policiamento de tráfego é o módulo do sistema responsável por garantir que aplicações mal comportadas não ultrapassem suas reservas previamente estabelecidas;
- Mecanismo de estabelecimento de reservas: Determina os mecanismos para estabelecer e manter uma determinada reserva de recursos para um fluxo. O protocolo RSVP foi criado com este objetivo.

3.1 Modelo de Serviços Integrados

Esse modelo é caracterizado pela reserva de recursos, de certa forma se caracterizando como um serviço orientado à conexão. Nas aplicações de tempo real, antes da transmissão dos dados, o caminho a ser utilizado deve ser definido, bem como os recursos a serem utilizados devem ser reservados. Para isso ser possível, um protocolo de sinalização denominado RSVP (Resource Reservation Protocol) é utilizado para configurar caminhos e reservar recursos, possuindo as características adequadas para servir como protocolo de suporte para o IntServ.

3.1.1 Protocolo RSVP

Trata-se do primeiro padrão industrial de protocolo para garantir QoS em redes heterogêneas. O RSVP é um protocolo de sinalização que têm a capacidade de requisitar um determinado nível de QoS através da rede. Ele carrega o pedido pela rede, visitando cada nó que a rede usa para carregar o fluxo. Este protocolo trabalha em conjunção com os protocolos de roteamento, utilizando sua robustez, além de pressupor algum tipo de política de escalonamento (Weighted Fair Queing - WFQ -

ou Random Early Detection -RED) implementada nos nós intermediários para implementar a reserva de recursos. O daemon RSVP, para efetuar a reserva de recursos num roteador, comunica-se com os módulos locais de controle de admissão e controle de política (policy control). O controle de admissão determina se o nó tem capacidade para fornecer a QoS requisitada. Já o controle de política determina se o usuário tem permissão administrativa para efetuar tal requisição. O RSVP deve ser implementado nó a nó, e, portanto, enfrenta alguns problemas de escalabilidade.

O protocolo RSVP apresenta os seguintes atributos [RFC 2205]:

- Faz reserva de recursos para aplicações unicast e multicast .
- RSVP é um protocolo simplex, isto é, faz reserva para fluxo unidirecional.
- RSVP é orientado a recepção (receiver-oriented) ; é a recepção que inicializa e mantém a reserva de recursos ao longo da rede para um determinado fluxo.
- Mantém um soft state em roteadores para prover suporte em casos de mudanças automáticas de rotas e alteração dinâmica do número de participantes. Através deste mecanismo o RSVP executa um refresh para manter as reservas asseguradas nos roteadores.
- Não é um protocolo de roteamento, mas depende diretamente dos protocolos que determinam as rotas.
- RSVP transporta e mantém o controle de tráfego e parâmetros de policiamento.
- RSVP é suportado pelo IPv4 e IPv6.
- O protocolo RSVP faz a reserva sobre um fluxo de dados, definido entre dois endereços IP - origem e destino, ou no caso de uma transmissão multicast com serviços diferentes para mais de um destino. Para o RSVP, é a recepção que solicita a reserva de recursos à rede.

3.1.2 Caracterização de Tráfego e Parâmetros de Controle

O transmissor irá se comprometer a gerar tráfego dentro de um determinado conjunto de regras. Os parâmetros aceitos pelo modelo de serviços integrados do IP são especificados pela RFC 2215. A origem do tráfego promete à rede que o tráfego por ela gerado terá um determinado perfil, ou seja, será conforme a um certo conjunto de regras. A rede utilizará esta informação para decidir se aceita o tráfego e, em caso positivo, como gerenciá-lo. Este perfil é utilizado para policiar o tráfego, ou seja, verificar se a origem está gerando tráfego de acordo com o que foi prometido.

O parâmetro utilizado para declarar a taxa de geração de tráfego é chamado de TOKEN.BUCKET_TSPEC, e é especificado de acordo com o modelo do algoritmo do balde de fichas (ver anotações em aula).

A aplicação irá se comprometer a gerar tráfego de acordo com os parâmetros declarados na reserva. Estes parâmetros são utilizados apenas para descrever as características do tráfego gerado.

Cada roteador será responsável por seu próprio mecanismo de policiamento de tráfego, o algoritmo do balde de fichas é utilizado como modelo porque seus parâmetros são comuns a vários algoritmos de policiamento de tráfego. Na seção seguir serão apresentadas duas caracterizações simples de fluxos de tráfego, que vêm sendo utilizadas atualmente em redes. A primeira, chamada de balde furado, é apropriada para os casos de transmissão num canal de capacidade constante, no sentido que limita a taxa máxima de transmissão de um fluxo de dados. A segunda, denominada balde de fichas , é apropriada para tráfego em rajada, limitando a sua taxa média.

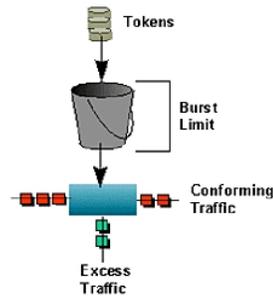


Figura 4: Balde de fichas

3.1.3 Balde furado

O objetivo deste algoritmo é converter um fluxo irregular numa seqüência regular de pacotes de tamanho fixo.

Cada fluxo é associado a seu próprio balde furado. Quando é preciso transmitir dados, os pacotes são inseridos no balde pelo transmissor. Estes pacotes serão retirados do fundo do balde e enviados pela rede com uma taxa r determinada pelo regulador no fundo do balde. O tamanho do balde, b , limita o tamanho da fila de espera para a transmissão pela rede se o balde já estiver cheio, os pacotes excedentes serão descartados. O fluxo de transmissão é regulado a uma determinada taxa, com intervalos entre pacotes de $\frac{d}{r}$ unidades de tempo (caso os pacotes sejam do mesmo tamanho, com d bytes).

3.1.4 Balde de Fichas

O balde de fichas é uma derivação do balde furado, utilizado para descrever fluxos com geração de tráfego variável. No caso do balde furado, um fluxo com taxa variável seria obrigado a especificar um perfil correspondente a sua taxa máxima, o que pode levar a um uso ineficiente da rede.

O algoritmo do balde furado será usado para gerenciar um dispositivo que regule o fluxo. Fichas serão colocadas no balde, de tamanho b , a uma taxa r . Se o balde ficar cheio, novas fichas são descartadas.

Quando um pacote deve ser transmitido, ele é colocado em um buffer. Para transmitir o pacote, o regulador deve retirar do balde o número de fichas igual ao tamanho do pacote, conforme indicado na Figura 4.

Suponha um pacote de tamanho d . Caso o número de fichas no balde seja maior que o d , o regulador irá remover d fichas do balde e o pacote será transmitido. Caso contrário, supondo que hajam b fichas no balde, o pacote deve esperar a inserção de $d - B$ fichas antes de ser enviado (ou o pacote pode ser descartado).

A utilização do algoritmo do balde de fichas permite irregularidade (rajadas) no fluxo, mas a limita. O balde furado garante que o fluxo nunca enviará mais do que r pacotes por unidade de tempo e o balde de fichas garante que, em um intervalo i , o fluxo não enviará mais do que $\frac{b+i}{r}$ fichas de dados, e que a taxa média de transmissão não excederá r . Uma outra diferença é que o balde de fichas nunca descarta pacotes: é atribuição do fluxo determinar o que fazer se não houver recursos da rede

suficientes para atendê-lo.

3.1.5 Relacionamento entre o RSVP e o CBQ

O CBQ e o RSVP podem trabalhar em conjunto: os objetivos de cada um são complementares. O RSVP deve ser capaz de configurar os parâmetros para o CBQ, que é parte do Kernel do sistema operacional.

Os modos de reserva suportados pelo RSVP são compatíveis com a capacidade do CBQ. Considerações devem ser realizadas quanto à limitação do WRR como algoritmo de escalonamento, é esperado um certo nível de erro com uma pequena flutuação no ajuste da taxa reservada é particularmente prejudicial com o serviço garantido (aumento do parâmetro D). Este nível de flutuação também depende da velocidade do processador utilizado e do tamanho dos pacotes.

É possível configurar as classes independentemente do RSVP, utilizando uma tabela de rotas modificada. Porém, o RSVP proporciona maior flexibilidade e dinamismo na utilização da rede conforme as necessidades do usuário.

3.1.6 Serviço de Carga Controlada

O serviço de carga controlada oferece ao cliente um fluxo com qualidade de serviço equivalente ao que ele receberia caso a rede não estivesse congestionada, e é definido pela RFC 2211. Este serviço não garante aspectos como a variação do retardo de transmissão, apenas realiza uma distribuição justa da banda passante aos diversos fluxos, utilizando um compartilhamento estatístico do enlace. O comportamento da reserva será descrita por parâmetros do balde de fichas.

Os parâmetros utilizados são a taxa média de transmissão (r , dado em bytes por segundo, é a taxa de inserção de fichas no balde), o tamanho do balde (b , dado em bytes, é o tamanho do balde de fichas), uma taxa de pico (p , bytes por segundo), o tamanho mínimo da unidade policiada (m , dado em bytes) e o tamanho máximo do pacote (M , dado em bytes).

O policiamento de tráfego é realizado da seguinte maneira: a quantidade de dados enviados não deve exceder $r.T + b$, onde T é o um período de tempo. Todos os pacotes de tamanho inferior a m devem ser computados como sendo de tamanho m . Pacotes que chegam a um elemento e causem violação no limite $r.T + b$ são considerados não conformes. Adicionalmente, pacotes de tamanho maior que M também são considerados não conformes. Segundo a RFC 2211, a taxa de pico (parâmetro p), apesar de ser considerado importante para alguns algoritmos de escalonamento de pacotes, pode ser ignorado para o serviço de carga controlada. A especificação para utilização deste parâmetro no serviço de carga controlada pode ser realizada formalmente no futuro. Todo o tráfego que excede o limite $r.T + b$ deve ser enviado em regime de melhor esforço, se houver recurso disponível.

3.1.7 Serviço Garantido

A RFC 2212 estabelece o padrão para o serviço garantido. O serviço garantido (o retardo máximo na transmissão do pacote deve ser determinístico) permite prover circuitos que garantem à aplicação uma parcela da banda passante e o atraso de propagação. É um circuito equivalente a uma conexão ponto a ponto.

Para o serviço garantido, o roteador exige uma TSpec (especificação de tráfego) e uma Rspec (especificação de serviço) para o fluxo. A TSpec consiste em uma especificação de tamanho do balde de fichas (b), uma taxa de esvaziamento do balde (r), uma taxa de pico (p), o tamanho máximo do pacote (M) e a unidade mínima policiada (m), como definido no item 0. A Rspec consiste em uma banda reservada R e um tamanho de buffer B .

É presumido que o roteador irá calcular, da TSpec, o espaço em buffer B que o fluxo pode consumir dentro do roteador. R é dado em bytes por segundo e B em bytes. A definição do serviço garantido reside no fato de que o atraso no modelo de fluido para um fluxo obedecendo a uma taxa dada pelo par (r, b) e sendo servido por um enlace com uma banda R é limitado por b/R desde que R não seja menor do que r . O serviço é garantido com uma banda R .

O policiamento de tráfego para o serviço garantido é realizado da seguinte forma: durante o tempo de transmissão, a quantidade de dados transmitidos não poderá exceder $[M + \min(p.T, r.T + b - M)]$. Datagramas que excedem este limite devem ser considerados não conformes. Pacotes com um tamanho menor do que m devem ser considerados de tamanho igual a m , e os pacotes com um tamanho maior que M serão considerados não conformes. O elemento de rede deve garantir que o atraso no envio de qualquer datagrama seja menor que $b/R + C/R + D$, onde C e D descrevem o desvio máximo local do modelo de fluido. É importante enfatizar que C e D são máximos. Então, por exemplo, se uma implementação tem falhas ocasionais no serviço (talvez devido ao processamento da atualização de rotas), D precisa ser grande o suficiente para acumular o tempo que um datagrama pode perder durante a falha no serviço.

O termo de erro C é dependente da taxa reservada. Ele representa o atraso de um datagrama que o fluxo pode experimentar dentro dos seus parâmetros. Um exemplo deste erro é a necessidade de converter um datagrama em células ATM, com as células sendo enviadas a uma frequência de $1/r$. Este parâmetro é dividido por R devido à dependência da taxa reservada.

O termo de erro D é dependente apenas do roteador (ou elemento de rede) e representa o pior caso da variação de tempo deste elemento. Por exemplo, em algoritmos de escalonamento de pacotes, o algoritmo Round Robin destina slots de tempo a cada fluxo. O termo D representaria, neste caso, o tempo máximo que um fluxo, uma vez pronto para ser enviado, tem que esperar seu slot.

3.1.8 Sinalização para QoS

Sinalização é a troca de informações entre os componentes de uma rede (como a PSTN - Public Switched Telephone Network, ou Rede Comutada de Telefonia Pública) necessária para prover e manter o serviço de comunicação. No caso da PSTN (rede telefônica usual), a sinalização inclui o tom de discar, a digitação do número do telefone, o recebimento do tom de espera ou de ocupado, etc.

O RSVP - Resource reSerVation Protocol - é o primeiro padrão industrial de protocolo para garantir QoS em redes heterogêneas. Trata-se de protocolo de sinalização que têm a capacidade de requisitar um determinado nível de QoS, carregando o pedido através rede e visitando cada nó que esta usa para carregar o fluxo. As aplicações em tempo real utilizam o RSVP para reservar recursos nos roteadores ao longo dos caminhos da transmissão, de maneira a que a largura de banda necessária possa estar disponível quando a transmissão realmente ocorrer. **Visão Geral**

O RSVP é um protocolo que permite a comunicação de pedidos de reserva de qualidade de serviços para determinados fluxos IP. Desta maneira, os roteadores e hosts devem ser capazes de interpretar os pedidos de reserva, bem como manter e atualizar estas reservas quando necessário.

Suponha um elemento A (transmissor) que deseja iniciar uma conexão com um elemento B (receptor), com uma determinada qualidade de serviço. Inicialmente o elemento A irá enviar uma mensagem para B, informando a sua intenção de iniciar uma transmissão e os parâmetros de qualidade de serviço desejados.

Caso o elemento B possua condições para estabelecer a conexão (a quantidade de recursos disponíveis de B deve ser igual ou superior ao requisitado por A), B irá enviar uma mensagem para A contendo uma confirmação do pedido de estabelecimento da reserva. Esta mensagem de confirmação deverá ser interpretada por todos os roteadores do caminho de B até A, e cada um deles deverá efetuar a reserva e enviar adiante a mensagem. A mensagem de confirmação deverá percorrer o mesmo caminho percorrido pela mensagem de requisição da reserva.

Caso qualquer um dos roteadores do caminho não possua recursos suficientes para estabelecer a reserva, então a conexão não será estabelecida. Quando a mensagem enviada por B atingir A, a conexão será estabelecida. Os passos detalhados para estabelecer uma reserva entre dois elementos com uma distância entre eles de 3 roteadores são descritos a seguir:

1. O equipamento A declara sua intenção de iniciar uma transmissão com uma determinada qualidade de serviço;
2. O roteador 1 envia a mensagem adiante sem realizar nenhum tipo de reserva;
3. O roteador 2 envia a mensagem adiante sem realizar nenhum tipo de reserva;
4. O roteador 3 envia a mensagem adiante sem realizar nenhum tipo de reserva;
5. O equipamento B realiza a reserva e envia uma mensagem respondendo positivamente ao pedido de A;
6. O roteador 3 realiza a reserva de recursos e envia a mensagem para o roteador 2;
7. O roteador 2 realiza a reserva de recursos e envia a mensagem para o roteador 1;
8. O roteador 1 realiza a reserva de recursos e envia a mensagem para o equipamento A. A reserva está estabelecida e pode ser utilizada.
- 9.

Conceitos Básicos

O RSVP é utilizado por um host para requisitar qualidade de serviço para uma aplicação em particular. O RSVP também é utilizado por roteadores para entregar requisições de qualidade de serviço para todos os nós ao longo do caminho para atingir o destino, estabelecendo e mantendo a qualidade de serviço requisitada. Geralmente as requisições irão resultar em recursos sendo reservados em cada nó ao longo do caminho escolhido para atingir o destino.

As requisições de qualidade de serviço são realizadas em uma direção (simplex), o protocolo trata o transmissor como sendo logicamente distinto do receptor. Isso não impede que o processo transmissor seja tratado como transmissor e receptor ao mesmo tempo.

O RSVP opera sobre o protocolo UDP, utilizando na camada de rede o IP versão 4 ou versão 6. Também é importante enfatizar que este protocolo é independente de outros protocolos, como ICMP, IGMP ou protocolos de roteamento.

Os componentes do RSVP são:

1. Transmissores
2. Receptores
3. Roteadores

Um transmissor avisa ao receptor que deseja enviar dados e qual a qualidade de serviço é necessária. O receptor então envia uma mensagem aos roteadores no caminho para que eles reservem recursos para esta transmissão. Uma vez que todos os roteadores tenham reservado os recursos, o receptor avisa ao transmissor que ele pode iniciar a transmissão de dados. Caso algum componente não possua condições para reservar os recursos necessários, o roteador ou host irá recusar o pedido de reserva. Os dois conceitos chave do protocolo RSVP são o Fluxo e a Reserva.

Fluxos

O RSVP define uma sessão como sendo um fluxo de dados com um destino em particular e um protocolo de camada de transporte [5]. Um fluxo é identificado por um rótulo no cabeçalho do pacote contendo a tripla (DestAddress, ProtocolId, [DstPort]).

O campo ProtocolId contém a identificação do protocolo de transporte utilizado. O campo DestAddress indica o endereço de destino, que pode ser unicast ou multicast e o campo opcional DstPort indica qual a porta no TCP ou UDP ou um campo equivalente em um outro protocolo de transporte utilizado.

Antes de enviar um fluxo, o transmissor envia uma mensagem de solicitação de reserva, destinada ao receptor. A mensagem contém o endereço IP de origem, o endereço IP de destino e a banda passante desejada. A banda passante desejada, especificada através da taxa de transmissão e do atraso máximo para este fluxo é a qualidade de serviço que o fluxo está requisitando. Esta mensagem é então enviada ao receptor pelos hosts e roteadores ao longo do caminho. Quando houverem múltiplos destinos, a mensagem de solicitação de reserva será enviada para todos os destinatários. **Reserva**

Uma vez recebida a mensagem de solicitação de reserva, o receptor se torna responsável por realizar a reserva. Com o receptor fazendo a reserva, existe grande flexibilidade para manipulação de fluxos multicast. Este modelo baseado no receptor é mais eficiente em ambientes de rede heterogêneos.

O receptor envia uma mensagem de reserva alertando cada roteador ou host na rede a sua intenção de receber um fluxo de dados com uma determinada qualidade de serviço. Esta requisição inclui o endereço IP de origem, o endereço IP de destino, a taxa de transferência necessária e o

tempo de atraso mínimo. Durante o processo de reserva, o pedido de Qualidade de Serviço é analisado por dois módulos de decisão: o Controle de Admissão e o Policiamento de Admissão. O Controle de Admissão determina se o nó possui recursos disponíveis para fornecer a Qualidade de Serviço determinada. O Policiamento de Admissão determina se o usuário possui permissão administrativa para fazer a reserva. Se ambos os testes tiverem sucesso, os parâmetros são configurados na camada de enlace do roteador para reservar a Qualidade de Serviço requisitada. Se qualquer um dos testes falhar, a aplicação é notificada e nenhuma reserva é realizada.

Quando o transmissor obtiver a resposta, contendo o resultado do pedido de Qualidade de Serviço requisitada, ele poderá iniciar a transmissão do fluxo (caso a reserva tenha sucesso) ou não (caso não tenha sucesso).

O responsável pela manutenção da reserva é o receptor, enviando periodicamente mensagens de manutenção do fluxo. Caso contrário, os roteadores no caminho podem desfazer a alocação de recursos devido à falta de notícias dos hosts que utilizam os recursos. Neste caso é razoável presumir que a comunicação cessou e remover o estado de reserva.

A estrutura e conteúdo do pedido de Qualidade de Serviço são documentadas em especificação desenvolvida pelo Integrated Services Working Group. Em sumário, o RSVP possui os seguintes atributos:

1. RSVP faz reserva de recurso para aplicações unicast e multicast (muitos para muitos), adaptando dinamicamente a mudança de grupos e rotas;
2. RSVP é simplex, ele faz reservas utilizando fluxos de dados unidirecionais;
3. RSVP é iniciado pelo receptor, isto é, o receptor de um fluxo de dados inicia e mantém a reserva de recursos utilizada para este fluxo;
4. RSVP mantém o estado da reserva nos roteadores e hosts, fornecendo suporte para manutenção dinâmica de grupos e adaptação automática a mudanças de rotas;
5. RSVP não é um protocolo de roteamento, mas depende dos protocolos de roteamento existentes ou futuros;
6. RSVP transporta e mantém os parâmetros de controle de tráfego e o controle de policiamento, que são opacos para o RSVP;
7. RSVP provê muitos modelos de reserva (também chamado de *estilos* de reserva), para atender a uma variedade de aplicações;
8. RSVP suporta a operação transparente através de roteadores que não implementam RSVP;
9. RSVP suporta Ipv4 e Ipv6;
10. Suporta receptores heterogêneos.

Estilos de Reserva

Uma requisição de reserva inclui um conjunto de opções que são coletivamente chamadas de *estilos de reserva*. Os tipos de reserva suportados pelo RSVP são listados na Tabela 1.

Seleção do transmissor	Distinta	Compartilhada
Explícita	Estilo Filtro Fixo (FF)	Estilo Explícito Compartilhado (EC)
Curinga	Não definido	Estilo Filtro de Curinga (FC)

Tabela 1: Estilos de Reserva do RSVP

Uma opção de reserva se preocupa com o tratamento de pedidos de reserva para diferentes transmissores com a mesma sessão: estabelecer uma reserva distinta para cada transmissor, ou senão fazer uma única reserva que é compartilhada entre todos os pacotes dos transmissores selecionados.

Outra opção de reserva controla a seleção de transmissores: pode ser uma lista explícita de todos os transmissores ou um curinga que implicitamente seleciona todos os transmissores para a sessão. Em uma seleção explícita de transmissores, cada parâmetro de filtragem deve coincidir exatamente com o transmissor, enquanto em uma seleção realizada com um curinga, nenhum filtro é necessário.

Estilo Filtro de Curinga (FC) O estilo FC implica a opção: reserva compartilhada e seleção de curinga para transmissor. Então, o estilo FC cria uma única reserva compartilhada pelos fluxos de todos os transmissores. Esta reserva pode ser vista como uma espécie de *tubo* compartilhado, onde a capacidade é a maior das requisições de reserva entre todos os receptores, independentemente do número de transmissores o utilizando. O pedido de reserva estilo FC é propagado para todos os transmissores e automaticamente estendido a novos transmissores à medida que eles surgirem.

Simbolicamente, pode-se representar uma reserva no estilo FC por $FC(*Q)$, onde o asterisco representa uma seleção curinga de transmissor e Q representa o desempenho requisitado.

O policiamento de tráfego será feito sobre o maior parâmetro solicitado entre todas as reservas realizadas, oferecendo flexibilidade para o suporte de grupos de multicast. Por exemplo, um host solicita uma reserva em estilo filtro de curinga, solicitando um determinado recurso. Caso algum parâmetro seja menor do que a reserva corrente, a reserva será atualizada com este valor maior. O policiamento de tráfego será realizado sobre todas as fontes de tráfego em relação à maior reserva.

Estilo Filtro Fixo (FF) O estilo FF implica as opções: reservas distintas e seleção explícita de transmissor. Assim, este estilo de reserva requisita a criação de uma reserva distinta para pacotes de dados de um transmissor em particular, não os compartilhando com pacotes de nenhum outro transmissor para a mesma sessão. Pode-se representar a reserva FF por $FF(SQ)$, onde S é o transmissor selecionado e Q é o desempenho requisitado; este par forma o descritor do fluxo.

O RSVP permite múltiplas reservas FF requisitadas simultaneamente utilizando uma lista de descritores: $FF(S1Q1, S2Q2, \dots)$. A reserva total em um enlace para uma dada sessão é a soma de $Q1, Q2, Q3, \dots$ para todos os transmissores requisitantes [5]. Neste estilo de reserva apenas um transmissor será ligado a uma qualidade de serviço, por exemplo, o endereço 200.250.250.1 possui uma reserva de recurso nos roteadores ao longo do caminho até o destino. Apenas o tráfego gerado por esta estação será policiado.

Estilo Explícito Compartilhado (EC) O estilo EC implica as opções: reserva compartilhada e seleção explícita de transmissor. Assim, uma reserva em estilo EC cria uma única reserva comparti-

lhada pelos transmissores. Diferente do estilo FC, o estilo EC permite que o receptor explicitamente especifique um conjunto de transmissores para serem incluídos. Pode-se representar uma requisição EC contendo uma requisição de reserva Q e uma lista de transmissores S1, S2, ... por SE((S1,S2,...)Q).

Reservas compartilhadas, criadas pelos estilos FC e EC são apropriadas para aplicações multicast que possuem múltiplas fontes de dados que transmitem preferencialmente uma de cada vez. Um sinal de áudio é um exemplo de aplicação capaz de compartilhar reservas, desde que um número limitado de pessoas fale simultaneamente, cada receptor pode utilizar o estilo FC ou EC para requisitar duas vezes a banda passante requerida para um transmissor (que habilita a fala simultânea até um determinado limite). Por outro lado, o estilo FF, que cria reservas distintas para fluxos de diferentes transmissores, é apropriado para sinais de vídeo.

As regras do RSVP desabilitam a mistura de reserva compartilhada com reservas distintas, uma vez que estes modos são fundamentalmente incompatíveis. Elas também desabilitam a mistura de uma seleção de transmissor explícito com uma seleção de curinga, desde que isto poderá produzir reservas não previsíveis para um receptor que especifique uma seleção explícita. Como resultado destas proibições, os estilos de reserva FC, FF e EC são mutuamente incompatíveis [5].

Suponha, por exemplo, uma lista de transmissores, com endereços IP 200.10.20.1, 200.10.20.2, 200.10.30.3 gerando tráfego na rede IP. Este tráfego será policiado de acordo com uma reserva de recursos única. Mecanismo de Funcionamento do Protocolo.

Tipos de Mensagem

A Figura 3 ilustra o modelo RSVP de um nó roteador. Cada fluxo de dados chega de um salto anterior através de uma interface de entrada e é colocado em uma ou mais interfaces de saída. A mesma interface pode assumir o papel de entrada e saída para fluxos de dados diferentes dentro da mesma sessão.

Existem dois tipos de mensagem fundamentais para o RSVP: Resv e Path. Cada host receptor envia mensagens de reserva (Resv) de volta para os transmissores. Estas mensagens devem seguir exatamente o caminho inverso que os pacotes irão utilizar, retornando para todos os hosts transmissores incluídos na seleção de transmissores. Eles criam e mantêm um *estado da reserva* em cada nó ao longo do caminho. As mensagens de reserva devem finalmente ser entregues para o host transmissor, que será capaz de configurar os parâmetros de tráfego apropriados para o primeiro salto.

Cada host transmissor envia uma mensagem de solicitação de reserva (Path) ao longo das rotas uni/multicast fornecidas pelo protocolo de roteamento. A mensagem Path armazena o *estado da rota* (path state) em cada nó ao longo do caminho. Este estado da rota inclui pelo menos o endereço unicast IP do nó correspondente ao salto anterior, que é utilizado para rotear a mensagem Resv salto-por-salto na direção reversa. No futuro, alguns protocolos de roteamento poderão vir a suprir esta informação, eliminando a necessidade da mensagem Path armazenar o caminho. Além desta informação, a mensagem Path contém as seguintes informações:

- Modelo de Transmissão: Uma mensagem Path é necessária para transportar uma descrição do transmissor, que descreve o formato dos pacotes de dados que o transmissor irá gerar. Esta

descrição está na forma de um filtro de desempenho que pode ser utilizado para seleccionar os pacotes deste transmissor de outros na mesma sessão dentro do mesmo enlace;

- Desempenho de Transmissão: Uma mensagem Path é necessária para transportar uma especificação de desempenho denominada TSpec, que define as características de tráfego do fluxo de dados que o transmissor irá gerar. A TSpec é utilizada pelo controle de tráfego para prever sobre-reservas, e falhas desnecessárias de controle de admissão;
- Adspec: Uma mensagem Path pode transportar um pacote de aviso OPWA (One Pass With Advertising [22]), conhecida como Adspec. Uma Adspec recebida em uma mensagem Path é enviada ao controle de tráfego local, o qual retorna uma Adspec atualizada; a versão atualizada é então enviada ao próximo nó do caminho. O objeto Adspec é modificado pelos elementos de rede para refletir as características do caminho (por exemplo, MTU), enquanto a mensagem de reserva está se propagando até o receptor.

Como as mensagens Path são enviadas com o mesmo endereço de origem e destino que o dado, então elas poderão ser roteadas corretamente através de nuvens não RSVP. Por outro lado, mensagens Resv são enviadas salto-por-salto; cada nó ativo RSVP envia adiante as mensagens Resv para o endereço unicast do salto RSVP anterior. No instante do estabelecimento da reserva, será inserida na mensagem um objeto do tipo ResvConf, contendo o endereço IP do transmissor, que será utilizado para enviar a confirmação do estabelecimento da reserva.

Remoção de Reserva

As mensagens de remoção de reserva removem o caminho ou o estado de reserva imediatamente. Apesar de não ser necessário remover uma reserva explicitamente, é recomendado que todos os hosts enviem uma requisição de remoção de reserva quando a aplicação termine e deixe de utilizar os recursos.

Existem dois tipos de mensagem de remoção de reserva: PathTear e ResvTear. Uma mensagem PathTear viaja através de todos os receptores para o transmissor e apaga o path state, bem como todos os estados de reserva dependentes ao longo do caminho. Uma mensagem ResvTear apaga o estado de reserva e viaja dos transmissores para os receptores, a partir de seu ponto de início. Uma mensagem PathTear (ResvTear) pode ser descrita como sendo o inverso de uma mensagem Path (respectivamente ResvTear). Uma requisição de remoção de reserva pode ser iniciada por uma aplicação em um sistema fim (transmissor ou receptor), ou por um roteador como resultado do temporizador de limpeza. Uma vez iniciada, a requisição de remoção de reserva deve ser enviada salto por salto o mais rapidamente possível. A mensagem causa a remoção dos recursos alocados em cada nó que a recebe. Como nos casos anteriores, esta mensagem será enviada imediatamente ao próximo nó, mas somente se houver alteração após a mudança. Como resultado, uma mensagem ResvTear irá retornar o estado da reserva de volta (somente) tão rápido quanto possível.

Como todas as outras mensagens RSVP, a requisição de remoção de reserva não possui garantia de entrega. A perda de uma mensagem deste tipo não irá causar uma falha no protocolo porque a reserva não utilizada irá eventualmente ser apagada através do mecanismo acionado pelo temporizador de limpeza ao invés de ser explicitamente destruído. Caso uma mensagem de remoção de reserva seja perdida em um determinado enlace, depois de um certo tempo, o roteador seguinte irá iniciar novamente o processo de remoção de reserva para os roteadores seguintes.

Deverá ser possível remover qualquer subconjunto de um estado estabelecido. Para o path state, a granularidade da remoção é um único transmissor. Para o estado de reserva, a granularidade é um filtro de desempenho individual. Uma mensagem ResvTear especifica o estilo e os filtros, qualquer desempenho do fluxo é ignorado. Qualquer desempenho de fluxo em operação será removido se todos os seus filtros de desempenho forem removidos.

Mensagens de Erro

Existem dois tipos de mensagem de erro: ResvErr e PathErr. As mensagens PathErr são muito simples, elas são enviadas para o transmissor aonde o erro se originou, e não alteram o path state nos nós que ela atravessa. Por exemplo, uma mensagem de erro pode ser gerada quando um nó não possui recursos suficientes para estabelecer uma reserva. Um nó também pode decidir interromper uma reserva estabelecida.

A interpretação de uma mensagem ResvErr é bastante complexa, pois um erro pode ser resultado de um pedido de reserva que agrupa vários fluxos e várias características de tráfego.

Desde que uma falha pode ser resultado de uma combinação de um número de requisições, um erro de reserva deve ser relatado para todos os receptores responsáveis. Em adição, combinar requisições heterogêneas cria uma dificuldade potencial conhecida como *problema do matador de reservas*, no qual uma requisição poderia negar o serviço para outras. Existem na verdade dois problemas deste tipo:

1. O primeiro problema (KR-I, Killer Reservation Problem) ocorre quando já existe um pedido de reserva Q0 em curso. Se outro receptor fizer uma reserva maior Q1 \geq Q0, o resultado da combinação Q0 e Q1 pode ser rejeitada pelo controle de admissão em algum nó. Isso não deve negar a requisição para Q0. A solução para este problema é simples: quando o controle de admissão falhar para uma determinada requisição de reserva, qualquer pedido de reserva existente é preservado.
2. O segundo problema (KR-II) é o seguinte: o receptor fazendo uma reserva Q1 é persistente mesmo quando o Controle de Admissão de algum nó nega o pedido de reserva. Isto não deve bloquear um receptor diferente de estabelecer uma reserva menor Q0 que teria sucesso se não fosse adicionada com Q1.

Para resolver este problema, uma mensagem ResvErr estabelece um estado adicional, chamado *estado bloqueado*, em cada nó por onde ela passa. O estado bloqueado em um nó modifica o procedimento de combinação para omitir o sobre-carregamento de desempenho de fluxo (Q1, no exemplo), permitindo que uma requisição menor seja enviada e estabelecida. O estado de reserva Q1 é dito *bloqueado*.

Uma requisição de reserva que é negada pelo Controle de Admissão cria um estado bloqueado, mas é deixado nos nós para trás do ponto de falha. Tem sido sugerido que estas reservas já realizadas até o ponto de falha representem reservas *descartáveis* e devam ser desativadas por temporização. Contudo, as reservas para trás são deixadas, pelas seguintes razões:

Existem duas possíveis razões para um receptor persistir em uma reserva que falha: (1) está realizando um pooling por um recurso disponível ao longo do caminho, ou (2) ele quer obter a QoS desejada ao longo de tantos caminhos quantos forem possíveis. Certamente no segundo caso, e talvez no primeiro, o receptor irá querer manter as reservas que foram feitas desde o ponto de falha;

Se esta reserva para trás não for mantida, a resposta do RSVP para certos estados transientes de falha será imprevisível. Por exemplo, suponha que uma rota seja alterada para uma rota alternativa que está congestionada, portanto uma reserva existente falha na alocação dos recursos nesta nova rota. Então ela facilmente poderá recuperar a rota original. O estado bloqueado em cada roteador não deve remover o estado de reserva ou bloquear a sua imediata atualização. **Mensagens Path**

Uma mensagem tipo PATH trafega periodicamente do transmissor ao receptor pelo mesmo caminho utilizado pelo fluxo de dados. Ela contém um objeto SENDER_TEMPLATE definindo o formato dos pacotes de dados e um objeto SENDER_TSPEC especificando as características de tráfego do fluxo. Opcionalmente, ela pode conter um objeto ADSPEC carregando dados de aviso (OPWA) para o fluxo. **Mensagens Resv**

Uma mensagem RESV carrega requisições de reserva dos receptores aos transmissores, percorrendo o caminho inverso do fluxo de dados para uma determinada sessão. O endereço IP de destino de uma mensagem RESV é o endereço unicast do salto anterior, obtido do estado de reserva. O endereço de origem deve ser o endereço do nó que originou a mensagem.