

2. O Estado da Arte em Qualidade de Serviço

Neste capítulo será apresentada uma visão geral das arquiteturas e protocolos que, atualmente, têm sido estudados e propostos na literatura. Também será abordada uma proposta já existente de mapeamento entre as arquiteturas de Serviços Integrados e Serviços Diferenciados.

2.1 Soluções de QoS Analisadas

2.1.1 Serviços Integrados (IntServ)

A arquitetura de Serviços Integrados [RFC1633] foi proposta a partir da necessidade de se viabilizar aplicações de tempo real na Internet. Até então tais aplicações eram impraticáveis em função de requisitos como alta sensibilidade a *jitter* e latência, atrasos nas filas e conseqüentes perdas nos congestionamentos, etc.

Um princípio básico da arquitetura de Serviços Integrados é o de implementar componentes baseados na extensão e não na modificação do serviço IP. O termo Serviços Integrados é usado para caracterizar um modelo na Internet que inclui o serviço de melhor esforço (*Best Effort*), o serviço de tempo real e o serviço de compartilhamento controlado do *link*.

A arquitetura de Serviços Integrados faz uso da premissa de que garantias não podem ser obtidas sem reservas. Alguns argumentos apresentados contra a reserva de recursos e contestados na [RFC1633] foram:

- a) A largura de banda será infinita. Contestada com o argumento de que isso jamais ocorrerá, uma vez que novas aplicações que demandam mais banda surgirão naturalmente com o aumento da infra-estrutura.
- b) A simples adoção de prioridade de encaminhamento é suficiente. Contestada com o argumento de que funciona apenas algumas vezes e sob certas condições. Quando surgirem muitos fluxos de tempo real, todos acabam sendo degradados.

- c) Aplicações podem se adaptar. Contestada com o argumento de que, mesmo aplicações com capacidade de adaptação, não eliminam a necessidade do pacote ser entregue em um tempo mínimo.

O modelo de Serviços Integrados é composto por quatro componentes: o classificador, o controle de admissão, o escalonador de pacotes e o protocolo de configuração de reserva (ver Figura 2.01).

O escalonador de pacotes coordena o encaminhamento de diferentes fluxos de pacotes usando um conjunto de mecanismos de filas e temporizadores. Há um outro componente que pode ser considerado parte do escalonador de pacotes chamado "*estimator*". Trata-se de um algoritmo com propriedades de medição que estima o uso do *link*.

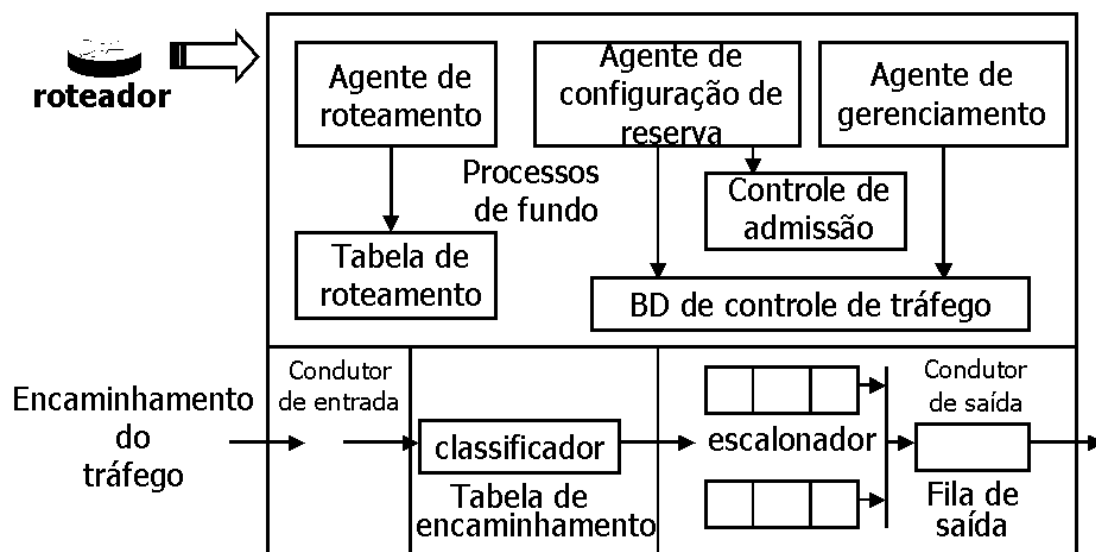


Figura 2.01 - Roteador IntServ

O classificador faz o mapeamento dos pacotes nas diversas classes existentes. Uma classe é uma abstração que pode ser local para o roteador, ou seja, o mesmo pacote pode ser classificado de formas diferentes por roteadores diferentes. Poder-se-ia, por exemplo, definir como pertencentes a uma certa classe todos os fluxos de vídeo ou ainda todos os fluxos atribuídos

a uma organização especial. Os pacotes da mesma classe teriam o mesmo tratamento no escalonador de pacotes.

O controle de admissão implementa, em *hosts* e roteadores, o algoritmo que decide se um novo fluxo, que requer QoS, deve ou não ser aceito. Isso ocorre a partir de uma avaliação no impacto a ser causado nas reservas já garantidas.

O protocolo de configuração de reserva, que é implementado pelos vários mecanismos mostrados na figura 2.01, é necessário para criar e manter o estado de cada fluxo nos *hosts* destinos e nos roteadores ao longo do caminho.

2.1.2 Protocolo de Reserva de Recursos (RSVP)

O RSVP (Resource Reservation Protocol) é um protocolo fim-a-fim compatível com o TCP/IP que provê suporte de QoS para aplicações através da reserva de recursos na rede. Além de suportar métodos de interconectividade, a infra-estrutura das redes é preservada. Uma analogia ao processo a reserva de recursos pode ser estabelecida com faixas exclusivas para trânsito de ônibus (ver Figura 2.03). Neste caso, por mais que haja congestionamento nas vias convencionais o trânsito flui normalmente nas áreas de reserva.

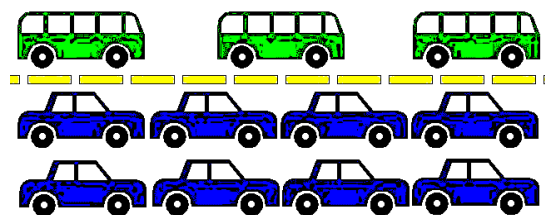


Figura 2.02 - Tráfego sem reserva

Figura 2.03 - Tráfego com reserva

O RSVP é um protocolo de controle comparável ao ICMP (Internet Control Management Protocol) ou protocolo de controle de mensagem na Internet. Ele foi projetado para operar com roteamentos unicast e multicast, ou seja, tanto para aplicações que são configuradas para um único receptor, quanto para as que têm o potencial de transmitir para mais de um receptor sem precisar enviar para a rede inteira (broadcast).

Com o RSVP as aplicações estão habilitadas a prover uma sinalização com os recursos que irão precisar. Para garantir a reserva, os *hosts* e roteadores afetados se comprometem em prover esses recursos. Se um dos roteadores não é capaz de provê-los ou os recursos não estão disponíveis, o *host* ou roteador pode recusar a reserva. A aplicação é notificada, imediatamente, de que a rede não pode suportá-la, evitando assim perda de tempo e dinheiro em tentativa e erro uma vez que os recursos da rede não serão consumidos caso a reserva não se concretize.

Quando um segmento torna-se comum, em uma árvore *multicast*, apenas uma reserva é efetivada com base nas requisições feitas aos transmissores por todos os receptores. Com isso, nesses pontos comuns, a transmissão sofre um processo de agregação, evitando reservas redundantes.

Se o transmissor fosse o responsável pelas reservas, ele precisaria saber as características de todos os possíveis receptores. Todavia, se essa tarefa couber ao receptor, ele precisará conhecer somente suas capacidades.

Quando o RSVP solicita uma reserva originada na extremidade do receptor, o atual controle da QoS ocorre na extremidade do transmissor. As requisições são passadas em um fluxo reverso do caminho de dados a partir do receptor para todos os transmissores. Em cada nó intermediário, duas decisões são tomadas:

- a) Fazer a reserva: A solicitação é passada ao nó pelo módulo de controle de admissão ou *admission control* e controle de política ou *policy control*. Se eles não permitirem que a reserva seja feita, uma mensagem de erro é enviada para o receptor(es) apropriado. Se a solicitação for aceita, os parâmetros da QoS desejada serão configurados no escalonador de pacotes e filtrados no classificador de pacotes pelas especificações de fluxo (*flowspecs*) e de filtro (*filterspecs*) respectivamente. (Figura 2.04)

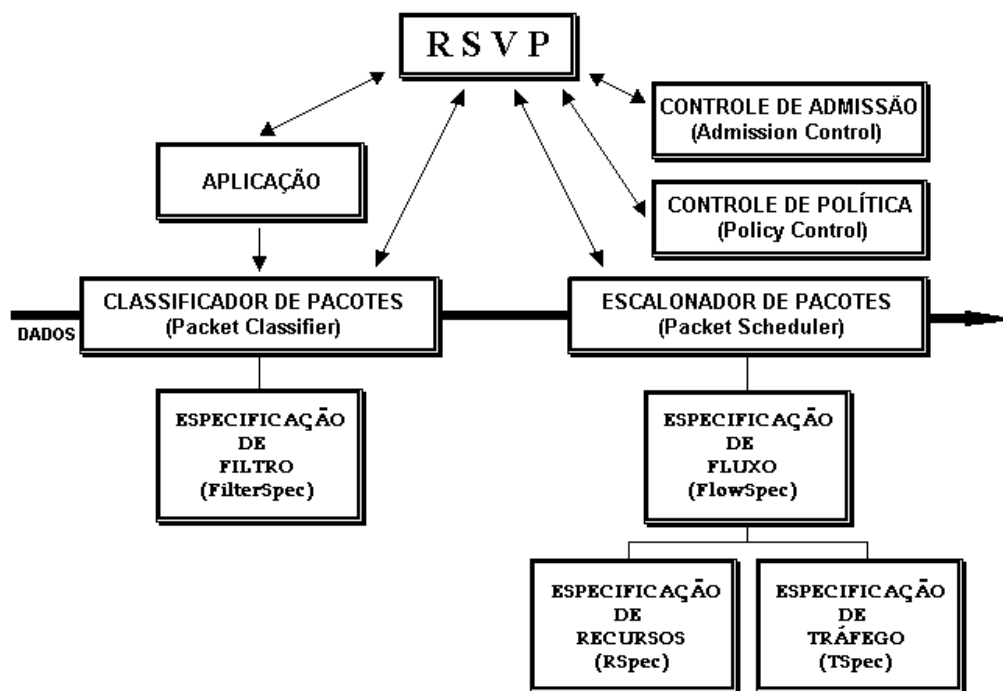


Figura 2.04 - Módulos RSVP

- b) Remeter a reserva especificada com o fluxo em direção aos transmissores apropriados. Contudo, esta requisição pode diferir em relação a que o nó recebeu do antecessor. Em particular, o nó pode modificar a especificação de fluxo (*flowspec*), ou ele pode sempre casar reservas a partir de vários *downstreams* receptores em árvores *multicast*, enviando somente as solicitações que, de alguma maneira, permitam que o máximo de especificações de fluxo (*flowspecs*) sejam recebidas (Figura 2.05).

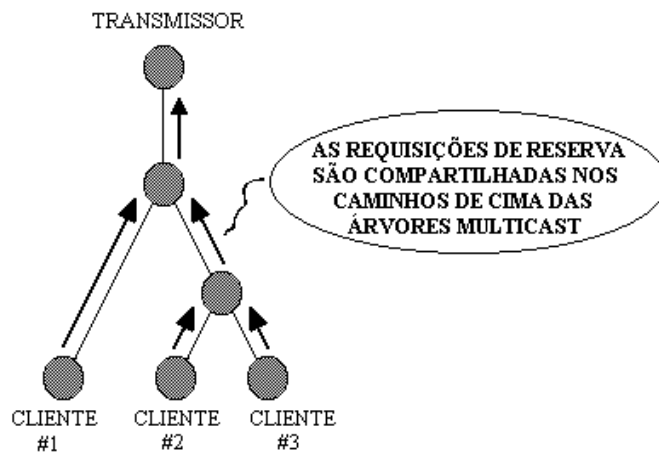


Figura 2.05 - Agregação de Reservas

O RSVP também provê um mecanismo chamado *one pass with advertising* que, usado pelo receptor, pode apresentar a QoS fim-a-fim que está sendo obtida. Esses anúncios podem ser usados pelo receptor para construir ou para ajustar as requisições de reserva.

A Figura 2.06 serve para ilustrar um processo de negociação de reservas entre o transmissor e o receptor. Nesse caso os seguintes passos serão executados [Xia99]:

- a) um transmissor inicia uma reserva RSVP com uma mensagem sobre caminho para um ou mais receptores. A mensagem "*PATH*" especifica as características do tráfego;
- b) cada roteador intermediário, ao longo do caminho, passa adiante a mensagem "*PATH*" de acordo com o que foi determinado pelo protocolo de roteamento;
- c) uma aplicação, no receptor, recebe a mensagem "*PATH*" e responde com uma outra mensagem que requisita recursos para o fluxo;

- d) o receptor envia uma mensagem de reserva "RESV" para roteadores ao longo do caminho;
- e) os roteadores checam a mensagem "RESV" e, se todas as condições forem satisfeitas, reservam os recursos;
- f) uma aplicação no transmissor recebe uma mensagem de reserva efetivada;
- g) o transmissor inicia o envio dos pacotes de dados.

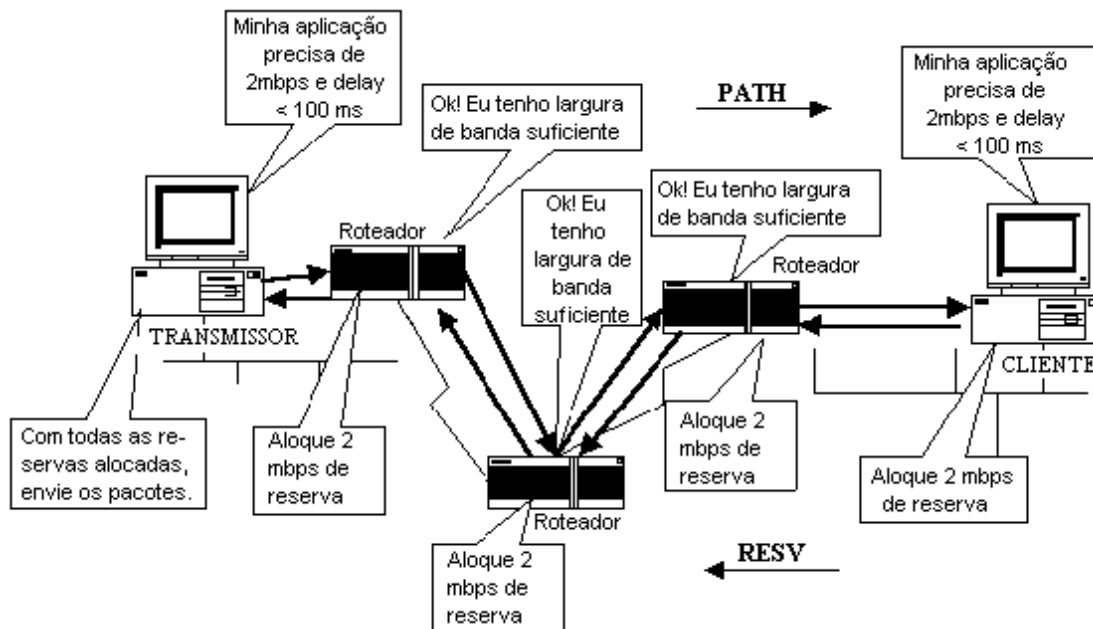


Figura 2.06 - Processo de Negociação de Reservas RSVP

Um dos principais problemas com o RSVP é o número de controles gerados em tabelas mantidas nos roteadores. Como cada fluxo (vídeo, áudio, etc.) requer um controle independente, feito através do assinalamento de parâmetros nas tabelas dos diversos roteadores da sub-rede, torna-se inviável a reserva em função da sobrecarga desses roteadores, posto que, o número de fluxos pode vir a se tornar extremamente grande. Isso empobrece as propriedades escalares do protocolo [Hus98] e, portanto, torna-se o

principal entrave para a utilização dessa tecnologia. Outro problema é a sobrecarga causada pelo excesso de mensagens de sinalização.

2.1.3 Serviços Diferenciados (DiffServ)

A arquitetura de Serviços Diferenciados (DiffServ) prima por adotar mecanismos que mantêm o modelo escalar mesmo com o crescimento significativo do número de *host* na Internet. Além disso nenhum processo de sinalização é requerido eliminando os problemas causados pela sobrecarga de mensagens. Basicamente a arquitetura se utiliza de diferentes agregações de fluxos de acordo com o perfil das aplicações ou dos usuários.

Os pacotes IP são marcados com diferentes prioridades pelo usuário ou pelo Provedor de Serviços Internet (ISP). De acordo com as diferentes prioridades das classes, os roteadores reservam o compartilhamento de recursos necessários (em particular largura de banda). Esta concepção habilita um provedor de serviços a oferecer diferentes classes de QoS por diferentes custos para seus usuários.

Para a marcação dos pacotes (ver figura 2.07) o chamado DS byte (byte de Serviços Diferenciados) é usado no cabeçalho de cada pacote IP. No IPv4 há um mapeamento do octeto *Type of Service* (ToS) e no IPv6 do *Traffic Class* (TC). Seis bits desse byte, chamados *Codepoint*, são combinados para definir o comportamento do pacote por salto ou PHB (*Per Hop Behavior*) que é analisado em cada roteador no despacho do pacote. Os outros dois bits foram preservados para uso em futuras propostas, são os chamados CU (*Current Unused*). [Fen99]

O significado dos bits individuais nos PHBs ainda não foram padronizados, entretanto, uma série de propostas já estão sendo analisadas no grupo de discussão Diffserv da IETF [WGDS99].

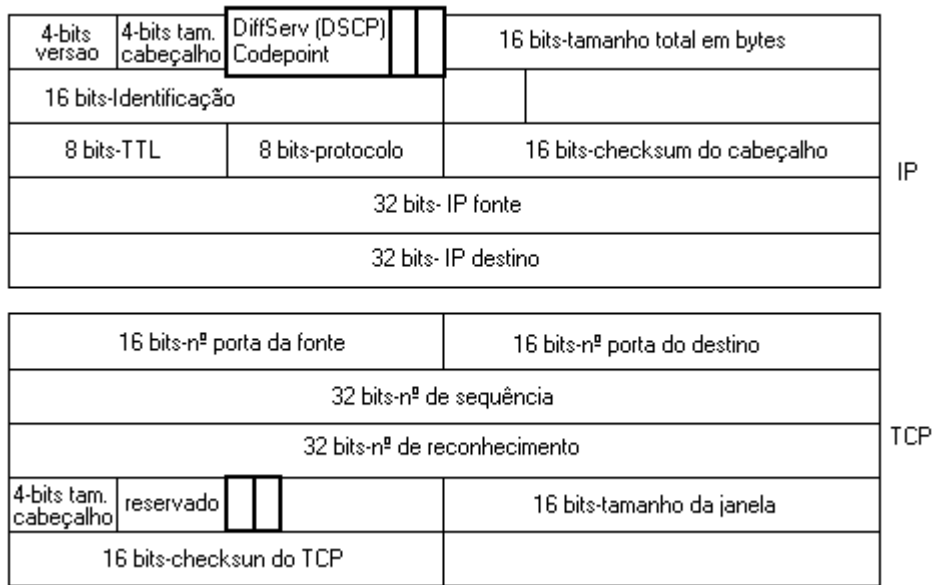


Figura 2.07 - Lay-out do IP/TCP

Duas propostas de serviços de encaminhamento nos roteadores merecem destaque, são elas: Serviço Assegurado ou AF (**A**ssured **F**orwarding) e Serviço Expresso ou EF (**E**xpedited **F**orwarding).

2.1.3.1 O PHB AF

O PHB AF procura fornecer um domínio de Serviços Diferenciados (DS) para oferecer diferentes níveis de despacho para pacotes IP. Para isso, quatro classes AF são definidas e ficam hospedadas em cada um dos nós DS reservando um certo montante de recursos de despacho (espaço de *buffers* e largura de banda). Os pacotes IP que desejarem usar os serviços providos pelos PHBs AF são marcados pelo usuário ou Provedor de domínio DS dentro de uma ou mais classes de AF de acordo com os serviços que o receptor houver definido.

Para cada classe AF os pacotes IP são marcados com um de três valores que indicam a probabilidade de descarte (Tabela 2.01). Em caso de

congestionamento, o descarte será feito considerando-se a classe e a probabilidade de descarte escolhidos. O nó congestionado irá tentar proteger os pacotes com baixa prioridade de descarte.

Em um nó DS, o nível de despacho assegurado em um pacote IP depende, basicamente, de três fatores: da quantidade de recursos de despacho que foram reservados para a classe AF que o pacote pertence; da carga das classes AF e da probabilidade de descarte dos pacotes.

Um nó DS **não pode** reordenar pacotes AF do mesmo microfluxo quando eles pertencerem a mesma classe AF. Se isso ocorresse a integridade da aplicação estaria comprometida.

Não há requisitos quantitativos de tempo (atraso e *jitter*) associados com o encaminhamento de pacotes AF. O controle de tráfego é feito pelos nós de borda. Quando necessário eles podem suavizar o tráfego, descartar pacotes, aumentar ou diminuir a probabilidade de descarte e remarcar pacotes para outras classes AF.

Uma série de códigos denominados *CODEPOINTS* são sugeridos (Tabela 2.01) para identificar as classes e prioridades de descarte: [RFC2597]

DESCARTE	CLASSE			
	Serviço Olímpico			Classe 4
	Classe 1/Ouro	Classe 2/Prata	Classe 3/Bronze	
Baixo	AF11 = 001010	AF21 = 010010	AF31 = 011010	AF41 = 100010
Médio	AF12 = 001100	AF22 = 010100	AF32 = 011100	AF42 = 100100
Alto	AF13 = 001110	AF23 = 010110	AF33 = 011110	AF43 = 100110

Tabela 2.01 - Proposta de Codificação para o PHB AF

Um dos serviços sugeridos [RFC2597], chamado serviço olímpico, trabalha apenas com três classes: Ouro, Prata e Bronze e com três níveis de descarte: baixo, médio e alto. O serviço Ouro, com baixa prioridade de descarte (AF11), poderia ser usado para transmissão de algumas aplicações de multimídia, desde que sejam elásticas, ou seja, possam se adaptar a situações de congestionamentos. Isso pode ser conseguido, por exemplo, com o uso de *buffers*.

2.1.3.2 O PHB EF

O PHB EF pode ser usado para prover um serviço fim a fim, através de domínios DS, com baixos níveis de perda, baixa latência, baixo *jitter* e largura de banda assegurada. Como o serviço aparece nos pontos finais, como uma conexão ponto a ponto ou uma linha dedicada virtual, ele também é descrito em [Nich99] como serviço Premium.

O mecanismo EF é o mais adequado para encaminhamento de pacotes que fazem parte de fluxos multimídia, uma vez que certas garantias podem ser dadas o que não ocorre com o mecanismo AF.

Perdas, latência e *jitter* ocorrem devido ao tráfego nas filas dos roteadores enquanto os pacotes caminham na rede. Portanto, prover baixos níveis de perdas, latência e *jitter* para algum tráfego agregado, implica garantir que essa agregação não veja filas. Se isso ocorrer, elas devem ser bem pequenas. Criar um serviço EF requer basicamente dois passos:

- a) Configurar os nós para que as agregações tenham uma taxa mínima de despacho bem definida. Isso implica uma certa independência do estado dinâmico do nó, ou seja, uma independência da intensidade de outros tráfegos no nó.

- b) Condicionar as agregações, via policiamento e suavização, para que a taxa de chegada nos nós seja sempre menor que a taxa mínima de encaminhamento configurada no nó.

O primeiro item é provido pelos PHBs EF enquanto que o segundo pelos mecanismos de policiamento de tráfego localizados nos nós de borda. Se o PHB EF é implementado por um mecanismo que permite preempção ilimitada de outros tráfegos, ou seja, uma fila prioritária, a implementação precisa incluir algum mecanismo que limite o perigo do tráfego EF infligir em outro tráfego. Isso pode ser feito, por exemplo, através do limitador conhecido como balde de fichas (token bucket - Figura 2.08). O tráfego que exceder o limite deve ser descartado. Esta taxa máxima de EF e o tamanho das eventuais rajadas devem ser configurados pelo administrador da rede.

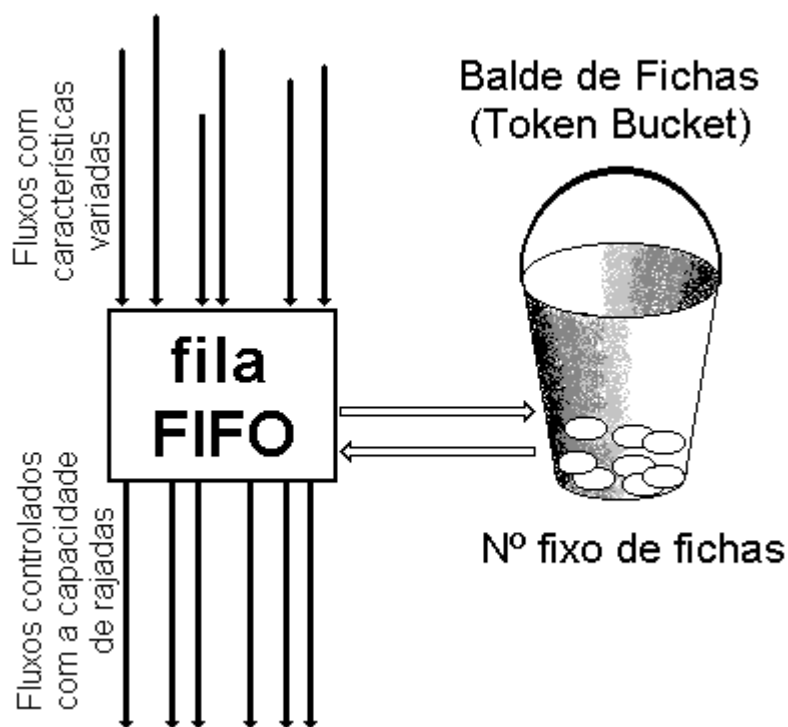


Figura 2.08 - Balde de Fichas

A analogia com o balde de fichas se apresenta da seguinte forma: Para que um fluxo passe adiante, ele precisa de um determinado montante de

créditos, representados pelas fichas do balde. Se o balde estiver vazio ou com um número de fichas inferior ao necessário, o fluxo terá que esperar até que novas fichas sejam depositadas. Com isso obtém-se um certo controle da vazão no domínio, ou seja, pode-se provisionar a rede configurando-se os diversos parâmetros do balde.

A implementação dos PHBs EF pode ser conseguida utilizando-se múltiplas filas com diferentes prioridades (ver Figura 2.09), o que pode ser feito, por exemplo, através de um escalonador *round robin*, com pesos.

O *Codepoint* recomendado para identificar um pacote de fluxo EF é '101110'. Os pacotes marcados como sendo de um fluxo EF não podem ser indicados para outros PHBs que não os EF. [RFC2598]

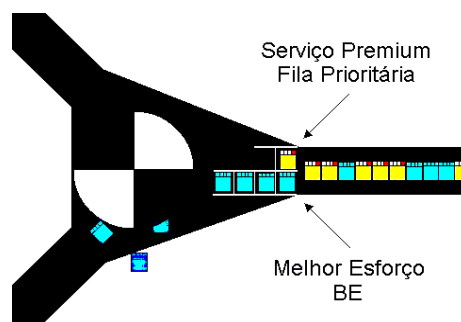


Figura 2.09 - PHB EF

2.1.3.3 Provisionamento

Para se obter eficiência no uso de Serviços Diferenciados é fundamental que se faça uma estimativa de uso dos recursos suportados pela rede e se adote políticas para garantir o cumprimento dos perfis de tráfego negociados com os usuários. Para isso uma série de mecanismos devem agir em conjunto. Os principais são [RFC2465]:

- a) Classificador (Classifier): uma entidade que seleciona pacotes com base em informações contidas nos cabeçalho de acordo com regras definidas.
- b) Descartador (Dropper): um dispositivo que executa descarte de pacotes.
- c) Marcador (Marker): um dispositivo que marca o DSCP dos pacotes.
- d) Medidor (Meter): um dispositivo que faz medição.
- e) Suavizador (Shaper): um dispositivo que suaviza a passagem de pacotes, mantendo o fluxo dentro do perfil contratado com o usuário. (Ex.: Token Bucket).
- f) Condicionador de Tráfego (Traffic Conditioner): é uma entidade que executa o condicionamento de tráfego. Um condicionador pode conter medidores, marcadores, descartadores e suavizadores. A ação a ser adotada dependerá da política escolhida pelos administradores do domínio.

2.1.4 Engenharia de Tráfego (TE)

Segundo [Xia99], arquiteturas como Serviços Integrados (RSVP) e Serviços Diferenciados provêm prioridades diferenciadas para alguns tráfegos quando a carga da rede está alta. Quando a carga está leve (*light*), Serviços Integrados, Serviços Diferenciados e Serviço de Melhor Esforço diferem pouco em termos de desempenho. Então por que não evitar o congestionamento ao invés de propor mecanismos para sobrepujá-lo? Esta é a motivação para Engenharia de Tráfego.

Os principais motivos para o congestionamento das redes são a falta de recursos de rede e a distribuição desigual do tráfego. No primeiro caso, todos os roteadores e enlaces estão sobrecarregados e a única solução é prover

mais recursos aumentando a infra-estrutura. No segundo caso, algumas partes da rede estão sobrecarregadas enquanto que outras estão com a carga leve. As desigualdades na distribuição de tráfego podem ser causadas por protocolos de roteamento dinâmico como RIP e OSPF porque eles sempre selecionam o menor caminho para enviar os pacotes. Como consequência, roteadores e enlaces ao longo do menor caminho entre dois nós podem vir a se tornar congestionados enquanto que roteadores e enlaces ao longo do maior caminho são subutilizados.

Engenharia de Tráfego, portanto, é o processo de se organizar como o tráfego flui através da rede de modo que o congestionamento causado pela utilização desigual da rede possa ser evitado.

A Engenharia de Tráfego pode ser implementada manualmente, em redes pequenas, ou através de técnicas automatizadas como MPLS e Roteamento com QoS.

2.1.5 Protocolo de Encaminhamento Baseado em Rótulos (MPLS)

Como, em uma rede, um pacote viaja de um roteador ao seguinte, cabe a cada roteador uma decisão independente no processo de encaminhamento desse pacote. Para isso, cada roteador analisa o cabeçalho do pacote e executa um algoritmo de distribuição na camada de rede. Escolhe-se, independentemente, o salto seguinte para o pacote baseado na análise do cabeçalho do pacote e nos resultados do algoritmo de distribuição. A questão é que o cabeçalho dos pacotes contém consideravelmente mais informações do que o necessário para escolher o salto seguinte e essa quantidade excessiva de informações pode consumir tempo de pesquisa desnecessariamente. Escolher o salto seguinte pode, conseqüentemente, ser pensado como uma composição de duas funções: A primeira divide o jogo inteiro de pacotes em uma série de classes de equivalência de encaminhamento (FECs - Forwarding Equivalence Classes). A segunda traça cada FEC a um salto seguinte. Assim que a decisão de encaminhamento é tomada, os pacotes traçados no mesmo FEC são tratados sem distinção. Todos os pacotes que pertencem a um FEC particular e que viajam de um nó específico seguirão o mesmo trajeto (ou se determinados tipos da distribuição *multi-path* estiverem em uso, todos seguirão um de uma série dos trajetos associados com o FEC).

Em MPLS (Multiprotocol Label Switching), a atribuição de um pacote a um FEC é feita apenas uma vez, porque o pacote se incorpora a rede. O FEC a que o pacote é atribuído é codificado com um valor fixo e curto através de uma "etiqueta". Quando um pacote é enviado a seu salto seguinte, a etiqueta será emitida junto com ele; isto é, os pacotes serão etiquetados antes que estejam enviados.

Em saltos subseqüentes, não há nenhuma análise adicional no cabeçalho da camada de rede do pacote. Isso agiliza bastante o processo de encaminhamento. Além disso, a etiqueta será usada enquanto houver um deslocamento predeterminado em uma tabela que especifique o salto seguinte. Quando necessário, a etiqueta velha é substituída por uma etiqueta nova, e o pacote é enviado a seu salto seguinte.

No paradigma do encaminhamento de MPLS [Ros99], uma vez que um pacote é atribuído a um FEC, nenhuma análise adicional do cabeçalho é feita por roteadores subseqüentes (ver Figura 2.10). Todo o encaminhamento é dirigido pelas etiquetas. Isto tem uma série de vantagens sobre o encaminhamento convencional da camada de rede. Uma delas é o fato do protocolo operar simulando o processo de encaminhamento do ATM com os campos definidos no nível de enlace [Dav98]

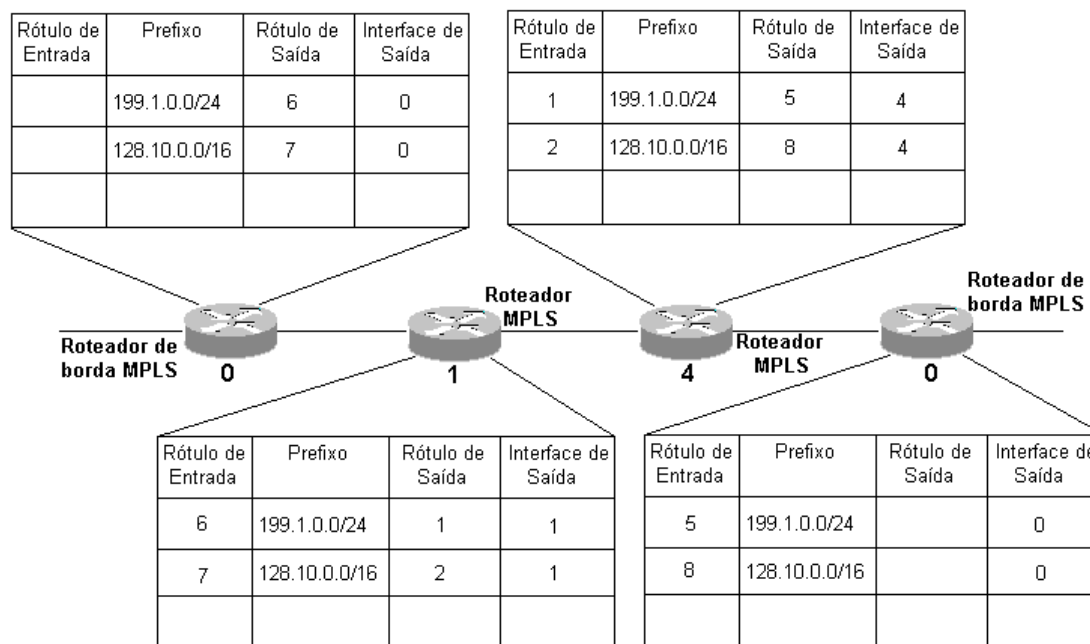


Figura 2.10 - MPLS

A Figura 2.10 mostra dois fluxos onde os endereços IP são "substituídos" por rótulos simples de entrada e saída. Essa técnica agiliza consideravelmente o

processo de encaminhamento de pacotes uma vez que não se faz necessário a pesquisa nas robustas tabelas de roteamento.

2.1.6 Roteamento Baseado em Restrições

O processo de roteamento atual considera como parâmetros de encaminhamento apenas o caminho mais curto (RIP - Routing Information Protocol) ou o peso administrativo de cada enlace (OSPF - Open Shortest Path First). Todavia, nem sempre esses requisitos refletem as necessidades das aplicações. Fazer um roteamento baseado em restrições de Qualidade de Serviço ou *Constraint Based Routing* implica em usar métricas como largura de banda, atraso e jitter no processo de decisão de encaminhamento dos pacotes. [RFC2386]

Quando se usa roteamento com QoS, os caminhos para os fluxos podem ser determinados com base no conhecimento de recursos disponíveis na rede e nos requisitos de QoS do fluxo.

Os principais objetivos do roteamento baseado em QoS são:

- a) Determinação dinâmica dos caminhos possíveis: consiste em determinar um caminho, a partir de muitas possibilidades de escolha, que tenha uma boa probabilidade de acomodar os requisitos de QoS do fluxo. A determinação desse caminho pode ser baseada em políticas como o custo do caminho, o provedor selecionado, o horário, etc.
- b) Otimização dos recursos usados: baseado no esquema de roteamento, pode-se utilizar, de forma mais eficiente, os recursos da rede, aumentando-se, conseqüentemente, a vazão total da rede.

2.2 Protocolo de Gerenciamento de Políticas (COPS)

O modo como a maioria dos dispositivos determinam se podem ou não permitir a reserva de recursos pode ser de duas formas: interno, verificando a possibilidade de reserva de largura de banda ou externo, verificando se o requisitante tem permissão para solicitar a reserva. O mecanismo que os dispositivos usam para obter informações externas é COPS (Common Open Policy Service), isto porque todos os dispositivos, ao longo do caminho, registram o estado dos fluxos do tráfego em um servidor COPS (Figura 2.11). Com isso, o servidor tem a habilidade de avaliar o impacto de um fluxo fim a fim. Cada roteador, por outro lado, vê apenas o comportamento por salto para cada fluxo individualmente, de forma simplificada, pode-se dizer que o COPS é um protocolo simples para obter informações especificamente relacionadas a políticas.

O COPS define um ponto de decisão política ou Policy Decision Point (PDP, normalmente referenciado como servidor de política) e um ponto onde se cumprem as políticas ou Policy Enforcement Point (PEP). O PEP e o PDP se comunicam através de conexões TCP, o PDP mantém o estado dos PEPs que têm requisitado informações de políticas e pode enviar atualizações, mesmo que não tenham sido solicitadas, se as informações de políticas mudarem. O PEP pode tomar decisões locais de admissão, entretanto, um módulo chamado Ponto de Decisão Local ou Local Decision Point (LDP) precisa, freqüentemente, informar ao PDP sobre suas decisões, assim como o PEP precisa deferir as mudanças impostas pelo PDP. Tendo pontos de decisão local e remoto, o sistema provê um mecanismo que temporariamente (quando o *link* com a autoridade central estiver caído por exemplo) pode tomar decisões de admissão.

A comunicação ocorre da seguinte forma:

- a) O PEP consulta o PDP através de uma porta TCP (normalmente 3288) para iniciar uma sessão de gerenciamento para políticas.
- b) O PEP pode ser configurado com o endereço IP do PDP correspondente ou ele pode confiar em outro mecanismo como o Protocolo de Serviço de Localização ou Service Location Protocol (SLP) para obter o endereço automaticamente.

A Figura 2.11 mostra um exemplo de negociação do COPS em uma rede RSVP Unicast [ALIS99] e a seguir serão apresentados os passos desta negociação:

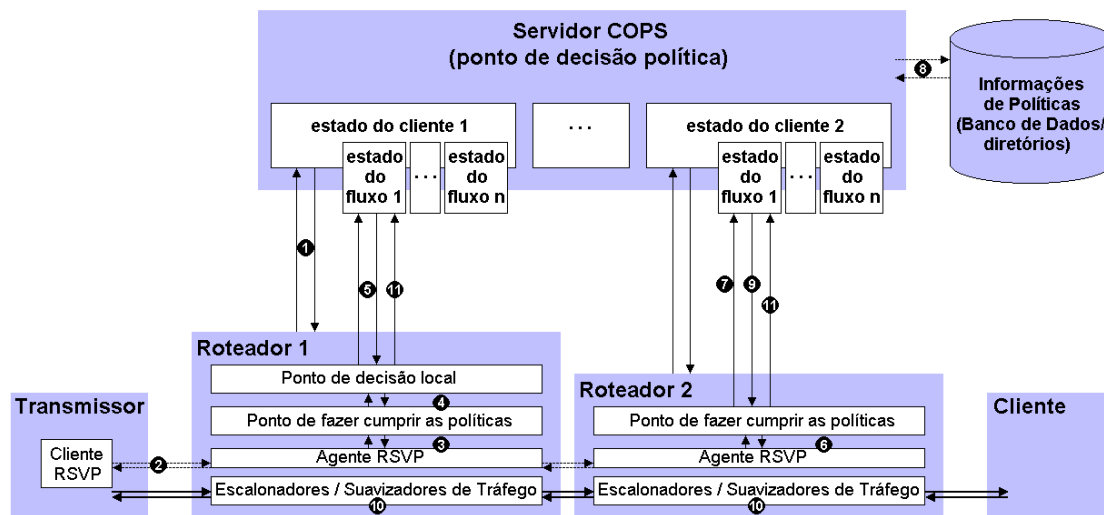


Figura 2.11 - Transações COPS em RSVP

- 1) O Roteador 1 se registra com o servidor de políticas, dizendo que políticas ele pode cumprir.
- 2) O Receptor RSVP requisita uma reserva de largura de banda ao longo do caminho.
- 3) O Roteador 1 consulta seu controle de admissão para ver se a requisição pode ser atendida.

- 4) O Roteador 1 (com LDP) toma uma decisão baseado nas políticas locais para permitir a reserva.
- 5) O LDP, então, registra o fluxo com o PDP, dizendo sua decisão, e mantém o estado.
- 6) O Roteador 2 (sem LDP) consulta o controle de admissão para ver se a requisição pode ser concedida.
- 7) O Roteador 2 registra o fluxo com o PDP e pergunta sobre a decisão política a ser tomada.
- 8) O servidor COPS checa as informações políticas (assim como as permissões do usuário) e toma a decisão.
- 9) O Roteador 2 recebe a decisão e permite a reserva.
- 10) O escalonador de tráfego aplica a reserva para os fluxos provenientes do transmissor, sujeito a mudanças pelo servidor COPS.
- 11) Quando o fluxo está atualizado, o PEPs informa ao servidor COPS.

2.3 Internet2

O projeto da Internet2 [I2-99] se solidificou em 1996, através da formação de um comitê geral de trabalho que reuniu 34 universidades norte-americanas. Pouco tempo depois, o governo do presidente Bill Clinton anunciou seu apoio a iniciativa e o interesse na criação e administração da NGI (*Next Generation Internet*) [NGI96]. A NGI é uma iniciativa norte-americana, voltada para o desenvolvimento de tecnologias e aplicações avançadas de redes Internet

para a comunidade acadêmica e de pesquisa. Em janeiro de 1997, mais de 100 universidades americanas já haviam assumido um compromisso formal em participar do projeto. Hoje, o projeto envolve 150 universidades norte-americanas, além de agências do governo e indústria e visa o desenvolvimento de novas aplicações como telemedicina, bibliotecas digitais, laboratórios virtuais, dentre outras que não são viáveis com a tecnologia adotada atualmente na Internet atual.

Em 1º de outubro de 1997 foi criada a UCAID (*University Corporation for Advanced Internet Development*). Trata-se de uma organização sem fins lucrativos cujo objetivo é orientar o avanço e desenvolvimento do projeto Internet2. Esta corporação, inicialmente constituída por três universidades americanas líderes no setor de pesquisa, tem como missão orientar os estudos e descobertas relativas às aplicações em todas as áreas do conhecimento, bem como em engenharia e ferramentas para redes eletrônicas de alto desempenho.

A UCAID dá uma organização formal às entidades que participam do projeto Internet2. Os líderes das 3 maiores universidades americanas especializadas em pesquisa de redes eletrônicas formam a atual diretoria da UCAID.

A arquitetura física da rede eletrônica que dá suporte à Internet2 inclui a implantação de GigaPOPs – pontos de presença com velocidade de tráfego da ordem de Gigabits[RNP2]. A função principal do GigaPOP é o gerenciamento da troca do tráfego Internet2 de acordo com especificações de velocidade e qualidade de serviços previamente estabelecidos através da rede. Cada GigaPOP tem a função de concentrar e administrar o tráfego de dados originados e destinados a um conjunto de universidades e centros de pesquisa localizados em uma mesma região geográfica.

Em seu estágio atual, a troca de dados entre os GigaPOPs é realizada por uma rede de alto desempenho mantida pela *National Science Foundation*, o VBNS (*Very High Performance Backbone Network System*) [VBN99]. A velocidade máxima oferecida pelo VBNS é de 622 Mbps, no entanto, a maioria das universidades que participam do projeto Internet2 opera com conexões de 155 Mbps em seus campus.

Em 14 de abril de 1998 foi lançado, nos Estados Unidos, o projeto de criação da maior e mais avançada rede norte-americana: a Abilene. O projeto foi desenvolvido pela UCAID e seu acesso feito através da rede nacional de fibras ópticas da Qwest, com tecnologia das empresas Cisco System e Nortel Telecom. A Rede Abilene representa um imenso potencial para o desenvolvimento de novas tecnologias Internet, oferecendo suporte para o desenvolvimento das aplicações que são o foco do projeto Internet2, tais como: laboratórios virtuais, bibliotecas digitais, ensino à distância, tele-medicina e tele-imersão, dentre outros. Esta rede oferece ainda a segurança, a confiabilidade e as inovações tecnológicas exigidas para a realização de pesquisas dentro do novo padrão I2 [I2-99].

Operando a 2,4 Gigabits por segundo, numa rede de mais de 16 mil quilômetros de fibra óptica, a Abilene oferece às instituições membros da UCAID oportunidades de interconexão num patamar de velocidade sem precedentes.

Um dos objetivos da Internet2 é viabilizar a realização de pesquisas e atividades de última geração em educação, fazendo uso da vantagem obtida com a "proximidade virtual" criada por uma infra-estrutura avançada de comunicação.

Assim, visando resolver problemas de acesso às informações, de compartilhamento e de uso de conteúdo educacional baseado na Internet, foi proposta a criação de infra-estrutura e de estratégias para suportar canais de

servidores (*server channels*) para a área acadêmica, o DSI (*Distributed Storage Infrastructure*).

Esses "canais", chamados Canais de Conteúdo Internet (*Internet Content Channels*), permitem que grupos de arquivos replicados sejam acessados pelos usuários finais de forma transparente. Isto é, embora os arquivos servidos por este projeto estejam replicados em múltiplos domínios e múltiplos servidores, um único URL é usado para acessar qualquer um deles, independentemente do domínio do servidor chamado para atender à requisição. Isso é conseguido através do uso de um servidor que direciona o pedido de informação para o servidor mais adequado.

A velocidade pode variar em função da quantidade de aplicações Internet2 que estejam sendo utilizadas nas redes conectadas ao GigaPOP. O ponto mais importante para o GigaPOP é que este deve ter certeza da capacidade adequada para antecipar a carga de tráfego. Espera-se um desempenho entre frações de uma conexão T3/DS3 (45 Mbps) até conexões OC-12 (*Optical Carrier 12* à velocidade de 622 Mbps).

Estão sendo utilizados enlaces dedicados (PVC - *Private Virtual Circuits*) ATM para conexão ao VBNS (*Very High Speed Backbone Service*), além de outros enlaces SONET (*Synchronous Optical Network*) [SON99]. Os enlaces entre os roteadores conectados por redes de longa distância serão fornecidos tipicamente por comutadores (switches) baseados em quadro (*frame-based*) ou em célula (*cell-based*), dependendo das necessidades de cada GigaPOP.

Os custos para a conectividade inter-GigaPOPs ainda não são conhecidos. Eles poderão variar em função das circunstâncias e dos serviços oferecidos. Portanto, é extremamente importante que os GigaPOPs guardem as estatísticas necessárias para uma alocação de custos razoável entre os membros do consórcio. Alguns objetivos são claros, entre eles:

- a) O custo de um serviço precisa ser previsível;
- b) Serviços de mais alto nível deverão custar mais que serviços de menor nível;
- c) A contabilidade deverá ser a mais simples possível, de modo a não consumir recursos de processamento apenas com contabilidade;
- d) No começo, o modelo a ser adotado será semelhante à Internet fase 1 com a divisão igual dos custos, talvez cobrando pela velocidade da conexão que, nesse caso, teria uma tarifa diferenciada .

Pode-se dizer, então, que o foco principal da Internet2 reside no desenvolvimento de aplicações avançadas com uso intensivo de tecnologias multimídia em tempo real.

Diversas aplicações já estão sendo desenvolvidas na Internet2, sendo que muitas delas se encontram em fase de teste. No momento, algumas das principais linhas de pesquisa desenvolvidas para a aplicação de serviços em redes de alto desempenho são:

- a) Bibliotecas digitais com capacidade de reprodução de imagens de áudio e vídeo de alta fidelidade; oferta de imagens de alta resolução com reprodução quase imediata na tela do computador e novas formas de visualização de imagens digitais;
- b) Ambientes colaborativos que englobam laboratórios virtuais com instrumentação remota; desenvolvimento de tecnologias para debates virtuais em tempo real, com utilização de recursos multimídia, em alta velocidade e de aplicação simplificada;

- c) Novas formas de trabalho em grupo, com desenvolvimento de tecnologias de presença virtual e colaboração em 3D;
- d) Telemedicina, incluindo diagnóstico e monitorização remota de pacientes;
- e) Projeção de telas de computadores em três dimensões, através da utilização da ImmersaDesk (espécie de grande tela de TV que projeta as imagens em 3D);
- f) Controle remoto de microscópios eletrônicos para pesquisas médicas.

2.3.1 Backbone com QoS (QBONE)

O projeto Qbone foi lançado na Universidade de Northwestern, em dezembro de 1998. Trata-se de uma iniciativa da Internet2 que visa fornecer sustentação a aplicações avançadas da rede Internet tais como voz, vídeo, teleimersão e Serviços Diferenciados (DiffServ) com Qualidade de Serviço (QoS) fim-a-fim. [RNP2]

Os participantes da iniciativa Qbone trabalham juntos para construir uma plataforma de testes internacional, cujo objetivo é o desenvolvimento e a realização de melhorias nas tecnologias emergentes de QoS. Estas tecnologias devem oferecer níveis garantidos de desempenho, superando alguns problemas atuais relacionados a congestionamento da rede.

2.4 Um modelo para operar redes de Serviços Integrados sobre Serviços Diferenciados

A IETF mantém um WG (Work Group) chamado ISSLL (Integrated Services over Specific Link Layers) [WGIS99] que discute propostas de mapeamento de Serviços Integrados sobre domínios de Serviços Diferenciados. No *draft* [Wro00] existe uma proposta de mapeamento sobre as classes AF e EF. Tal mapeamento prevê o controle tanto a nível de domínio quanto a nível do caminho percorrido pelo fluxo mapeado até o receptor (*PATH*). A informações necessárias ao mapeamento são obtidas via TSPEC (Traffic Specification). Para acomodar esse tráfego adicional, cria-se uma instância dos PHBs requeridos no mapeamento. No caso de mapeamento AF o controle será feito por um token bucket e no caso do mapeamento EF pelo escalonador. O mapeamento pode ocorrer tanto a nível de carga controlada (Controlled load Service), quanto de serviço garantido (Guaranteed Service).

O modelo consiste em usar a sinalização do RSVP para obter recursos tanto em domínios DiffServ como em domínios IntServ. Para viabilizá-lo, entretanto, faz-se necessário que a região DiffServ esteja apta a passar mensagens RSVP de forma que elas possam ser recuperadas na saída no domínio.

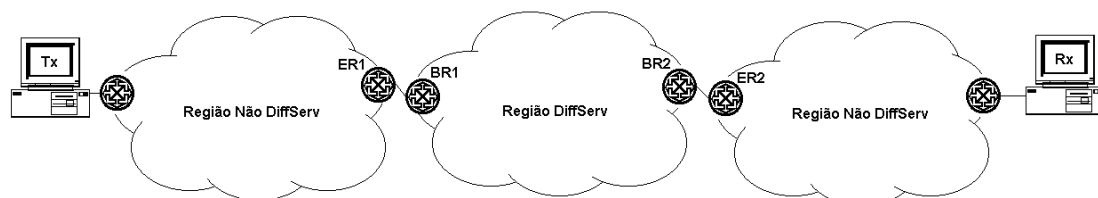


Figura 2.12 - Sinalização RSVP em Domínio DiffServ

Na Figura 2.12, assume-se que ambos os *hosts* Tx (Transmissor) e Rx (Receptor) fazem uso do protocolo de reserva de recursos (RSVP) para comunicar a quantidade de requisitos de QoS de suas aplicações. Deve-se

considerar a notação ER (Edge Routers) e BR (Border Routers) para referenciar os roteadores de borda.

A seguinte seqüência de passos ilustra o processo pelo qual uma aplicação obtém QoS fim-a-fim quando o RSVP é usado pelos hosts. [Ber99]

- a) O processamento de QoS no host de emissão Tx gera uma mensagem PATH do RSVP que descreve os requisitos de tráfego requeridos pela aplicação a ser transmitida.
- b) A mensagem PATH é carregada para o host de recepção, Rx. Na região da rede onde o transmissor é ligado, um processamento padrão de RSVP/Intserv é aplicado nos elementos de rede que o suportam.
- c) No roteador ER1 (da borda) a mensagem PATH é sujeita ao RSVP padrão que a processa e guarda seu estado PATH. A mensagem PATH é então emitida adiante (para a região da rede Diffserv).
- d) A mensagem PATH é ignorada por roteadores na região da rede de DiffServ e processada então em ER2 de acordo com regras de processamento padrão de RSVP.
- e) Quando a mensagem PATH alcança o host de recepção Rx, o sistema operacional gera uma mensagem de RSVP RESV, indicando o interesse em um serviço Intserv no tráfego oferecido.
- f) A mensagem RESV é então carregada para trás (para a região da rede de Diffserv e do host de emissão). Consistente com o processamento do RSVP/Intserv padrão, ela pode ser rejeitada por qualquer nó RSVP no trajeto se os recursos forem julgados insuficientes para atender o tráfego pedido.

- g) Em ER2, a mensagem RESV é sujeita ao processamento padrão do RSVP/Intserv. Pode portanto ser rejeitada se os recursos na relação downstream de ER2 forem julgados insuficientes para atender os recursos pedidos. Se não for rejeitada, será carregada de forma transparente através da região da rede de DiffServ, chegando em ER1.
- h) Em ER1, a mensagem RESV provoca processamento do controle da admissão. ER1 compara os recursos solicitados no pedido de RSVP/Intserv aos recursos disponíveis na região da rede de Diffserv (no nível de serviço correspondente de DiffServ). O nível de serviço correspondente é determinado pelo Intserv e a disponibilidade dos recursos é determinada pela capacidade de provisionamento no SLS (Service Level Specification). ER1 pode também aplicar uma decisão política tal que o pedido do recurso pode ser rejeitado baseado em critérios específicos da política do receptor, mesmo que os recursos agregados sejam determinados como disponíveis pelos SLS.
- i) Se ER1 aprovar o pedido, a mensagem de RESV será admitida e permitirá a continuidade para o remetente. Se rejeitar o pedido, o RESV não será enviado e as mensagens de erro apropriadas de RSVP serão emitidas. Se o pedido for aprovado, ER1 atualiza suas tabelas internas para indicar a redução da capacidade disponível no nível de serviço admitido.
- j) A mensagem RESV prossegue para a região da rede a qual o remetente é unido. Qualquer nó RSVP nesta região pode rejeitar o pedido de reserva devido aos recursos ou à política inadequada. Se o pedido não for rejeitado, a mensagem de RESV chegará no host de transmissão, Tx.

- k) Em Tx, o processamento de QoS recebe e interpreta a mensagem RESV recebida. Pode-se também ajustar um marcador apropriado de DSCP aplicado de acordo com a informação fornecida em RESV.
- l) O transmissor, Tx, pode também marcar o DSCP no cabeçalho dos pacotes que são transmitidos no fluxo de tráfego admitido. O DSCP pode ser escolhido com base no Intserv (tipo especificado na mensagem admitida RESV), ou pode ser um valor fornecido explicitamente em RESV.

Desta maneira, obtém-se QoS fim a fim como uma combinação das redes que suportam RSVP/Intserv e redes que suportam Diffserv.

2.5 Considerações Específicas

O modelo apresentado na seção 2.4, apesar de funcional, tem sua limitação por considerar apenas o mapeamento entre duas arquiteturas de QoS. No próximo capítulo será especificada a proposta de uma arquitetura "guarda-chuva", capaz de mapear não só IntServ e DiffServ, mas com abertura para que outras arquiteturas vindouras sejam incorporadas ao modelo.

