

Trabalho de Projeto de Redes de Computadores

Novembro/ 2001 - Professor: Oswaldo J.A. Franzin

Alunos:

Anna Paula Ferreira - ra 501006

Guilherme Steinberger Elias - ra 501022

Tema:

Moldagem de Tráfego

1. Introdução

Desde quando foi criado, o protocolo IP foi utilizado como um protocolo de comunicação onde o controle de tráfego utiliza apenas a regra do melhor esforço (Best-effort Service), ou seja, não provê nenhum mecanismo de priorização de tráfego baseado em serviços, portanto sem nenhuma garantia de alocação de recursos da rede. Na época da criação, não se imaginava que a Internet se tornaria a grande rede mundial atual e que haveria a necessidade de serviços como voz (telefonia), vídeo e dados em uma única infra-estrutura de redes de pacotes. Esta nova necessidade provocou uma corrida em busca de equipamentos, protocolos e algoritmos de redes que garantissem qualidade de serviços fim-a-fim.

O Traffic Shaping, aliado a Priorização de tráfego baseado em serviços e uso de novos algoritmos e protocolos de roteamento fazem parte da solução atual para esta necessidade, e serão vistos a seguir.

2. Para que serve a moldagem de tráfego?

Para evitar congestionamentos causados pelo tráfego de rajada, utilizando estratégias que fazem os hosts transmitirem em uma taxa uniforme e diminuírem a ocorrência de congestionamento. A moldagem de tráfego visa moldar o perfil de tráfego de acordo com as necessidades da rede e parâmetros contratuais.

Logo a moldagem de tráfego é uma atividade complementar ao policiamento de tráfego. Moldagem de tráfego é utilizada para "aplainar" o tráfego num perfil mais homogêneo, além de evitar a compactação de células o que provoca um aumento na variação do atraso. Obviamente procedimentos de moldagem de tráfego não devem permitir a violação de parâmetros de qualidade do serviço já assegurado às conexões.

O método de loop aberto, que ajuda a lidar com o congestionamento e força os pacotes a serem transmitidos em uma taxa mais previsível, é a estratégia de gerenciamento amplamente utilizada em redes ATM.

A esse gerenciamento de congestionamento chamamos de Moldagem de tráfego (traffic shapping).

3. Onde utilizar a moldagem de tráfego?

A moldagem de tráfego visa moldar o perfil de tráfego de acordo com as necessidades da rede e parâmetros contratuais.

Logo, tomando como exemplo a rede ATM temos as seguintes categorias de serviços:

CBR - Taxa de bit constante

rt-VBR - Taxa de bit variável para tempo real

nrt-VBR - Taxa de bit variável

ABR - Taxa de bit disponível

UBR - Taxa de bit não especificada

Características de cada categorias de serviços:

CBR:

- Garantia de largura de banda
- Serviço adequado para tráfego em tempo real

rt-VBR:

- Garantia de largura de banda
- Serviço adequado para tráfego em tempo real

nrt-VBR:

- Garantia de largura de banda
- Serviço adequado para tráfego em rajadas

ABR:

- Feedback sobre congestionamento
- Serviço adequado para tráfego em rajadas
- Garantia de largura de banda (Opcional)

UBR:

- Serviço adequado para tráfego em rajadas

Logo a moldagem de tráfego com seus algoritmos podem atuar efetivamente tanto na adequação para tráfego em rajadas quanto nos feedbacks sobre congestionamento, através dos pacotes de alertas.

Exemplos de onde se aplicam:

- O serviço nrt-VBR, se destina a aplicações sem restrições de tempo real mas que requeiram banda variável garantida e comunicação confiável (baixa perda de células).

Este serviço é normalmente utilizado na conexão à provedores de comunicação onde uma determinada banda contratada deve ser honrada pelo provedor.

- O serviço taxa de bit não especificado (UBR) se destina a aplicações sem necessidade de garantia de qualidade de serviço. O serviço UBR é um serviço do tipo "melhor esforço" e se destina a aplicações clássicas de rede tais como transferência de arquivos e correio eletrônico.
- O serviço com taxa de bit disponível (ABR) oferece às aplicações uma banda que varia em função da carga da rede, sendo um valor mínimo (eventualmente zero) e máximo estipulados pela aplicação e negociados com a rede. Um mecanismo de realimentação existe para o informe da banda disponível à aplicação. ABR se destina aplicações que possam tirar proveito da banda de rede assim que esta estiver disponível. Um exemplo de tal aplicação é a transferência de grandes massas de dados (imagens de satélites, por exemplo) do local onde são gerados, para um local remoto de armazenamento.

Ja utilizando o controle de congestionamento por Multicast é possível a moldagem de tráfego atuar também na largura da banda.

Exemplos de onde se aplicam:

- O serviço com taxa de bit constante (CBR) é definido para aplicações que requeiram uma banda (taxa de bit) constante garantida e baixo atraso na comunicação. Este serviço é conhecido como "emulação de circuito" pois emula um circuito comutado (tal qual uma conexão telefônica, por exemplo). CBR é utilizado principalmente na emulação de circuitos síncronos e na transferência em tempo real de áudio e vídeo não compactado.
- O serviço com taxa de bit variável para tempo real (rt-VBR) foi definido para aplicações com restrições de tempo real que requeiram banda variável garantida e baixo atraso na comunicação. rt-VBR é utilizado principalmente na transferência em tempo real de áudio e vídeo compactado.

4. Métodos

Os principais métodos utilizados para a moldagem de tráfego baseiam-se nos algoritmos de Balde Furado e de Balde de Tokens.

4.1 Balde Furado

Trata-se de um sistema de enfileiramento com um único servidor e que tem tempo de serviço constante. O host pode inserir um pacote a pulso de relógio (clock tick) na rede. Isso pode ser forçado pela placa da interface ou pelo sistema operacional. Esse mecanismo transforma um fluxo irregular de pacotes originados pelos processos de usuário dentro do host em um fluxo de pacotes regular para a rede, suavizando as rajadas e reduzindo muito a possibilidade de congestionamento.

4.2 Balde de Tokens

Nesse algoritmo, o balde furado retém tokens, gerados por um clock na faixa de um token a cada ΔT seg. O algoritmo de balde de tokens possibilita um tipo de moldagem de tráfego diferente do algoritmo de balde furado. O algoritmo de balde de tokens permite economia, até o tamanho máximo do balde, n . Essa propriedade significa que rajadas de até n pacotes podem ser enviadas simultaneamente, permitindo um certo volume no fluxo de saída e possibilitando uma resposta mais rápida a repentinas rajadas de entrada.

Outra diferença entre os dois algoritmos é que o de balde de tokens joga tokens fora quando o balde enche, mas nunca descarta pacotes. Em contrapartida o algoritmo de balde furado descarta pacotes quando cheio.

4.3 Conclusão sobre os algoritmos de Balde Furado e o Balde de Tokens

O algoritmo de Balde de Tokens e balde furado podem ser utilizados para amenizar o tráfego entre roteadores, assim como podem ser usados para controlar a saída do host. Entretanto, uma diferença óbvia é que um balde de tokens que está controlando um host pode fazê-lo interromper a transmissão quando necessário, porém mandar um roteador interromper a transmissão quando ela estiver em andamento pode resultar em perda de dados.

Logo, uma forma de obter um tráfego mais tranquilo é colocar um balde furado após o balde de tokens. A taxa do balde furado deve ser maior que o p do balde de tokens, mas deve ser inferior à taxa máxima de rede.

5. Especificação de Fluxo (loop aberto)

A moldagem de tráfego é mais efetiva quando o transmissor, receptor e a sub-rede concordam em relação a ela. Para que o acordo seja feito é necessário especificar o padrão de tráfego de um modo preciso. Um acordo como esse é chamado de especificação de fluxo (flow specification). Ele consiste em uma estrutura de dados que descreve o padrão de tráfego injetado tanto como a qualidade de serviço desejada pelas aplicações. As especificação de fluxo pode se aplicar aos pacotes enviados através de um circuito virtual ou também a uma seqüência de datagramas enviadas entre uma origem e um destino (ou vários destinos).

5.1 Variáveis que devem serem analisadas para o acordo

Características de entrada:

- tamanho máximo de pacote (bytes)
- taxa do balde de tokens (bytes/ s)
- tamanho do balde de tokens (bytes)
- taxa máxima de transmissão (bytes/ s)

Serviços desejados:

- sensibilidade a perdas (bytes)
- intervalo de perdas (us)

- sensibilidade a perdas em rajadas (pacotes)
- retardo mínimo percebido (*us*)
- variação de retardo máximo (*us*)
- qualidade de garantia

A sub-rede pode aceitar ou rejeitar a proposta ou voltar com uma contra-proposta com outras especificações de variáveis.

Uma vez que o transmissor e a rede tenham feito um acordo, o transmissor poderá perguntar ao receptor se ele também concorda.

5.2 Controle de Jitter

Utilizados para aplicações, que se importam com que o tempo de entrega dos pacotes seja constante. Logo o tempo de trânsito deve ser constante.

6. Traffic Shape

O traffic shape é um mecanismo disponibilizado pelos routers que permite ajustar a largura de banda para uma ou mais máquinas. Em serviços que requeiram grandes larguras de banda como por exemplo, transferência de ficheiros, áudio e vídeo é possível limitar e garantir uma largura de banda mínima para um mínimo de qualidade de serviço.

6.1 Moldagem genérica

Generic Traffic Shaping, (GTS) fornece mecanismos para controle de tráfego utilizando filtros conhecidos como 2-token bucket, limitando o tráfego de saída de uma interface a uma determinada taxa.

O tráfego classificado vai para um buffer limitador, sendo liberado sob regras pré-definidas de acordo uma política de controle de tráfego configurada pelo administrador ou derivada da interface.

Moldagem ou conformidade de tráfego pode ser útil em vários casos. Por exemplo, para limitar o tráfego de rajada de forma a não prejudicar o tráfego prioritário, reduzindo assim a latência ou em situações de congestionamento, para limitar um determinado tipo de tráfego não sensível a retardo, como transferência de arquivos, eliminando possíveis gargalos.

A moldagem GTS só se aplica em interfaces de saída, com o uso de listas de acesso para classificar e selecionar o tráfego. Funciona com qualquer tecnologia de enlace (nível 2) como Frame Relay, ATM (Asynchronous Transfer Mode), SMDS (Switched Multimegabit Data Service) e Ethernet.

O termo 2-token bucket é uma definição formal para taxa de transferência. Possui três componentes:

- um tamanho de rajada em bits (burst size), também chamado de committed burst (Bc) que especifica o quanto pode ser enviado num determinado intervalo de tempo;

- uma taxa média em bps, também chamada de CIR (Committed Information Rate), que especifica o quanto de dados, na média, pode ser enviado por unidade de tempo;
- um intervalo de tempo (T_c) em segundos, intervalo de medida que especifica o tempo por rajada.

6.2 Moldagem Frame Relay

Para controle e moldagem de tráfego em redes Frame Relay, utiliza-se o FRTS (Frame Relay Traffic Shaping), que oferece parâmetros úteis, como o CIR (Committed Information Rate), EIR (Excess Information Rate), FECN (Forward Explicit Congestion Notification), BECN (Backward Explicit Congestion Notification) e DE (Discard Eligible).

Com FRTS pode-se, por exemplo, limitar no CIR ou no EIR à taxa de pico do tráfego de saída em cada VC (Virtual Circuit), procedimento denominado rate enforcement. Pode-se obter uma maior granularidade no controle de tráfego, aplicando-se técnicas de enfileiramento como PQ (Priority Queueing) ou CQ (Custom Queueing) em cada VC ou no nível de subinterface.

Combinando-se CQ e enfileiramento por VC com rate enforcement, pode-se habilitar o VC a acomodar vários tipos de tráfego como IP, SNA e IPX, cada um com sua própria banda garantida.

7. Priorização de tráfego baseado em serviços

A procura por um meio de prover QoS em redes IP fez com que surgisse um conjunto de serviços os quais habilitassem um novo modelo de rede. Entre as técnicas de tratamento de tráfego já implementadas que se destinam a implementar serviços diferenciados eficientemente encontramos:

7.1 IP Precedence – para classificação de tráfego (edge)

O mecanismo de priorização IP Precedence prove a capacidade de particionar o tráfego dentro de múltiplas classes de serviços. O IP Precedence utiliza os três primeiros bits no campo Type-of-Service, no header IP, para especificar a classe de serviço para cada pacote. Com os 3 bits para o campo Precedência IP têm-se as seguintes configurações e funções:

- 0 - routine: estabelece a precedência rotina;
- 1 - priority: estabelece a precedência prioridade;
- 2 - immediate: estabelece a precedência prioridade;
- 3 - flash: estabelece a precedência flash;
- 4 - flash-override: estabelece a precedência flash-override;
- 5 - critical: estabelece a precedência crítica;
- 6 - internet: estabelece a precedência internetwork control;
- 7 - network: estabelece a precedência controle de rede.

Quanto maior o nível de classificação do pacote, maior será a prioridade no tratamento e alocação de recursos da rede. Os níveis 6 e 7 são reservados para as aplicações de controle e gerência da rede, ou seja, não é possível habilitar um pacote com 6 ou 7 e nem modificar um pacote já marcado. Todos os pacotes são normalmente marcados com o nível zero.

Nos roteadores Cisco, o campo precedência IP pode ser modificado através de listas de acesso (access lists) ou mapas de rotas (route maps). Nos gateways VoIP, podem ser habilitados em cada interface lógica dial peer voip.

A característica do IP Precedence habilita uma maior flexibilidade para tratamento de precedência. Esta característica prove a rede a capacidade de agir em modo passivo (onde a precedência é definida pelo usuário) ou ativo (onde o policiamento é definido pelo administrador da rede) baseado em IP, endereçamento MAC, porta física ou aplicação.

7.2 Committed Access Rate – para gerenciamento de largura de banda (edge)

O Committed Access Rate (CAR) tem como objetivo prover ao operador da rede a característica de gerenciamento da alocação da largura de banda a qual foi determinada na criação da conexão. O tratamento ocorre através de thresholds e pode ocorrer tanto em filas de entrada como em filas de saída dos roteadores. Como exemplo de CAR Policy podemos citar:

- Firm CAR – pacotes que excedem a banda alocada são descartados.
- CAR + Premium – pacotes que excedem a banda alocada são "recolored" com alta ou baixa preferência.
- CAR + Best Effort – pacotes que excedem a banda alocada são "recolored" até o estouro do threshold, depois são descartados.
- Per Application CAR – diferentes CARs são especificados para diferentes aplicações.

O CAR é o método para policiamento e controle de tráfego IP que realiza, basicamente, duas funções para qualidade de serviço:

- Gerenciamento de banda com limitação de taxa de acesso (policing) - permite controlar a taxa máxima de transmissão ou recepção de dados de uma determinada interface. O tráfego conformante é transmitido ou recebido, enquanto os pacotes que excedem os limites pré-definidos, ou são descartados, ou são reclassificados com outra prioridade para retransmissão;
- Classificação de pacotes através de precedência IP ou de grupos de QoS (um rótulo interno do roteador utilizado para definição de classes - permite particionar a rede em múltiplas classes de serviços ou níveis de priorização.

Para isso, utilizando também o mecanismo token bucket, o CAR examina o tráfego recebido na interface ou parte do tráfego selecionado pelos critérios das listas de acesso, compara a taxa de tráfego com a do token bucket e, de acordo com o resultado, toma uma ou várias ações. Ou seja, pode transmitir o pacote, descartá-lo ou reclassificá-lo com outro nível de prioridade, etc.

Os tokens são inseridos no balde na mesma taxa CIR. A profundidade do balde é o tamanho da rajada (burst size). Se houver tokens suficientes quando o tráfego chega ao balde, então o tráfego é dito estar em conformidade e a quantidade correspondente de tokens é removida. Se não houver tokens suficientes, então o tráfego é dito excessivo.

Os critérios de seleção de tráfego podem ser baseados em todo tráfego IP, em precedência IP, em listas de acesso (padrão ou estendida), em endereço MAC ou, ainda, em grupos de QoS. O endereço MAC e a precedência IP podem ser definidos através de listas de acesso rate-limit.

Com o CAR, pode-se limitar o tráfego por aplicação (Web, FTP, etc), por interface (p. ex., uma conexão serial a 2 Mbps, mas com acesso limitado em 512 Kbps), por endereço MAC (p. ex., controle de tráfego em PTT - Pontos de Troca de Tráfego), ou pode-se classificar ou reclassificar todo o tráfego de entrada num backbone a partir dos roteadores de borda, para tratamento diferenciado em termos de qualidade de serviços.

7.3 Weighted Fair Queueing – baseado em classe de serviços (backbone)

O Weighted Fair Queueing (WFQ) tem como finalidade proporcionar a capacidade de manipulação de tráfego de alta prioridade, requisitando baixo atraso, enquanto satisfatoriamente compartilha a banda restante entre tráfego de baixa prioridade. O WFQ divide o tráfego do link em fluxo de alta prioridade e baixa prioridade (baseado em métricas que incluem IP Precedence e volume de tráfego). Assim, o fluxo de alta prioridade recebe manipulação imediata enquanto fluxo de baixa prioridade é colocado junto e recebe uma proporção compartilhada da banda restante. Em certos casos, os pacotes de baixa prioridade podem ser descartados, durante períodos de congestionamento.

O algoritmo WFQ é uma implementação Cisco e seu algoritmo escalona o tráfego prioritário (interativo) para a frente da fila, reduzindo o tempo de resposta. Ao mesmo tempo, compartilha o restante da banda com os outros tipos de fluxo de uma forma justa. O WFQ é dinâmico e se adapta automaticamente às mudanças das condições de tráfego, sendo bastante útil em conexões seriais de baixa velocidade até 2 Mbps.

A fila WFQ já vem pré-configurada nas interfaces seriais dos roteadores Cisco. A classificação dos fluxos de dados pode ser realizada de diversas formas: endereço fonte ou destino, por protocolo, pelo campo precedência IP, pelo par porta/socket, etc. A quantidade de filas é configurável e a ponderação pode ser estabelecida por precedência IP, ou em conjunto com outros protocolos de QoS como o RSVP, ou ainda em tráfego Frame Relay, como VoFR (Voice over Frame Relay) por exemplo, através dos parâmetros FECN (Forward Explicit Congestion Notification), BECN (Backward Explicit Congestion Notification) e DE (Discard Eligible).

7.4 Tag Switching (backbone)

O tag switching é uma tecnologia de multilayer switching para backbone escalar de roteadores e switches incluindo Internet e Intranets corporativas. O tag switching combina capacidade de gerenciamento de tráfego e performance de nível 2 switching com a

escalabilidade e flexibilidade e nível de rede. O header incorpora informações de IP Precedence usado para indicar classes de serviços ou prioridade.

7.5 Resource Reservation Protocol (backbone)

O mecanismo Resource Reservation Protocol (RSVP) permite que aplicações requisitem uma Qualidade de Serviço específica. Hosts e roteadores usam o RSVP para entregarem seus requisitos para os roteadores ao longo do caminho do fluxo dos dados. O RSVP especifica que cada roteador entre os endpoints da rede participem de sessões de sinalização RSVP para realizar a reserva, teardown e gerenciamento dos recursos.

7.6 QoS em Redes Locais (802.1p)

Fala-se sempre em qualidade de serviços fim-a-fim, significando o caminho completo da origem ao destino do pacote, mas as tecnologias de QoS são mais comuns para redes de longa distância como Frame Relay, ATM, PPP, HDLC, etc. E, na rede local, o que fazer para que uma aplicação de voz sobre IP, como um telefone IP por exemplo, possa ter tratamento diferenciado da rede local até a rede de longa distância e daí chegar à rede remota com a qualidade que a aplicação requer?

A resposta está numa especificação de protocolo do IEEE denominada IEEE 802.1p priority queueing.

O 802.1p utiliza um procedimento similar ao de precedência IP quando define 8 níveis de prioridade de usuários (classes de tráfego), através de um rótulo (user_priority) de 3 bits que é transmitido no frame Ethernet.

Considerando, ainda, o exemplo do telefone IP, a precedência IP setada pelo telefone é mapeada no nível de prioridade 802.1p correspondente, e o frame Ethernet segue para o switch. Se o switch for compatível com 802.1p, ele classificará os frames Ethernet priorizando os de maior classe. Esse procedimento se repete até o frame chegar ao roteador que realiza o procedimento inverso, classificando os frames Ethernet e mapeando o nível de prioridade 802.1p na precedência IP correspondente. Na rede de longa distância, o pacote, já com a precedência definida, terá o tratamento de acordo com as técnicas de QoS já apresentadas. Se o switch não for compatível com 802.1p, ele simplesmente irá ignorar os rótulos e todos os frames serão transmitidos sem nenhum tratamento especial.

8. Controle e Inibição de Congestionamento

Há vários mecanismos de enfileiramento para controle e prevenção de congestionamento em interfaces de roteadores (Ethernet, seriais, Frame Relay, etc.) e switches nível 3, aplicáveis tanto em redes WAN como em LAN. As principais são apresentadas a seguir.

8.1 Random Early Detection – para gerenciamento de congestionamento (backbone)

A detecção RED - Random Early Detection (detecção randômica antecipada) é um mecanismo para prevenção e inibição de congestionamento ou congestion avoidance. O algoritmo monitora o tráfego antecipadamente utilizando as funções de controle de congestionamento TCP, descartando pacotes aleatoriamente e indicando para a fonte reduzir a taxa de transmissão, evitando assim situações de congestionamento antes que ocorra picos de tráfego. Quando habilitado numa interface, o RED começa a descartar pacotes a uma taxa que pode ser previamente configurada.

O Random Early Detection (RED) proporciona aos operadores da rede a habilidade de especificar facilmente a manipulação do policiamento de tráfego, de modo a maximizar o throughput em condições de congestionamento. O RED trabalha em conjunto com protocolos como TCP, de modo a poder se esquivar do congestionamento da rede através de algoritmos que possam:

- distinguir entre rajadas de tráfego temporários os quais podem ser acomodados pela rede,
- trabalhar cooperativamente com fontes de tráfego, para evitar inícios de baixas oscilações as quais podem criar ondas periódicas de congestionamento de rede,
- prover redução de tráfego satisfatório de modo a reduzir o tráfego fonte na proporção da largura de banda utilizada inicialmente. Assim, o RED trabalha com o TCP antecipando e gerenciando o congestionamento durante períodos de tráfego árduos, maximizando o throughput através de gerenciamento de perda de pacotes. Além, de proporcionar ao operador de rede a possibilidade de configurar queue depth thershold mínimo e máximo.

8.2 Weighted RED – para classes de serviço baseadas em gerenciamento de congestionamento (backbone)

É uma implementação da Cisco que combina as funcionalidades do RED com a classificação de pacotes por precedência IP. Baseado nessa classificação, o WRED descarta pacotes seletivamente, descartando inicialmente os pacotes de menor prioridade, com diferentes pesos para cada classe.

É possível desabilitar a classificação precedência IP e habilitar o descarte com base apenas no tamanho do buffer da fila; ou ainda utilizar o WRED em conjunto com o RSVP para se obter um descarte mais seletivo. Nesse caso, antes que ocorra uma situação de congestionamento, os fluxos de menor prioridade nas sessões RSVP serão descartados antes dos outros de maior prioridade.

O WRED é útil em qualquer interface na qual a possibilidade de congestionamento seja eminente. Entretanto, é geralmente utilizado em roteadores centrais de backbone (core routers), com a precedência IP habilitada pelos roteadores de acesso (edge routers).

O WRED combina a capacidade de IP Precedence e RED de modo a prover características diferenciadas de performance para diferentes classes de serviços. Assim, prove manipulação de tráfego preferencial para pacotes de mais alta prioridade. O WRED também prove manipulação preferencial de tráfego sobre condições de congestionamento sem agravar o próprio congestionamento.

Como Funciona a detecção RED

A detecção RED - Random Early Detection (detecção randômica antecipada) é um mecanismo para prevenção e inibição de congestionamento ou congestion avoidance. O algoritmo monitora o tráfego antecipadamente utilizando as funções de controle de congestionamento TCP, descartando pacotes aleatoriamente e indicando para a fonte reduzir a taxa de transmissão, evitando assim situações de congestionamento antes que ocorra picos de tráfego. Quando habilitado numa interface, o RED começa a descartar pacotes a uma taxa que pode ser previamente configurada.

WRED, ou Weighted RED, é uma implementação da Cisco que combina as funcionalidades do RED com a classificação de pacotes por precedência IP. Baseado nessa classificação, o WRED descarta pacotes seletivamente, descartando inicialmente os pacotes de menor prioridade, com diferentes pesos para cada classe.

É possível desabilitar a classificação precedência IP e habilitar o descarte com base apenas no tamanho do buffer da fila; ou ainda utilizar o WRED em conjunto com o RSVP para se obter um descarte mais seletivo. Nesse caso, antes que ocorra uma situação de congestionamento, os fluxos de menor prioridade nas sessões RSVP serão descartados antes dos outros de maior prioridade.

WRED é útil em qualquer interface na qual a possibilidade de congestionamento seja eminente. Entretanto, é geralmente utilizado em roteadores centrais de backbone (core routers), com a precedência IP habilitada pelos roteadores de acesso (edge routers).

8.3 FIFO - First In First Out

Em geral, o controle de tráfego nas conexões seriais dos roteadores é implementado através de filas FIFO (o primeiro a entrar é o primeiro a sair). Uma fila FIFO é um mecanismo de armazenamento e repasse (store and forward) que não implementa nenhum tipo de classificação.

A ordem de chegada dos pacotes é que determina a alocação da banda, e o que chega primeiro é logo atendido. É o tratamento default da fila nos roteadores, já que não requer nenhuma configuração. O problema ocorre em tráfego de rajada, que pode causar longos atrasos em aplicações sensíveis ao tempo. Por isso, filas FIFO não servem para aplicações que requerem QoS.

8.4 Enfileiramento Fair Queueing

No algoritmo de Enfileiramento Fair Queueing (enfileiramento justo), as mensagens são ordenadas em sessões, e, para cada sessão, é alocado um canal. A ordem na fila é realizada através do último bit que atravessa o canal. Essa operação provê uma alocação mais justa da banda entre os fluxos de dados.

8.5 Enfileiramento Priority Queueing

Numa fila com Enfileiramento Priority Queueing - PQ (enfileiramento prioritário), o tráfego de entrada é classificado em quatro níveis de prioridade: alta, média, normal e baixa (high, medium, normal e low). Os pacotes não classificados são marcados, por default, como prioridade normal.

Durante a transmissão, o tráfego classificado e marcado como prioritário tem preferência absoluta. Por isso, este método deve ser utilizado com cuidado, para evitar longos atrasos e aumento de jitter nas aplicações de menor prioridade. Num caso extremo, o tráfego de menor prioridade pode até nunca ser transmitido, se o de maior prioridade tomar toda a banda. Isso pode acontecer em conexões de baixa velocidade. Além disso, a fila default sempre tem que ser habilitada. Caso contrário, todo fluxo não classificado (sem uma correspondente lista de prioridade) também poderá não ser enviado.

Há várias opções de classificação de tráfego numa fila PQ. A classificação pode ser por protocolo (IP, IPX, DecNet, SNA, etc), por interface de entrada ou por lista de acesso.

Este método apresenta boa qualidade de sinal de voz no caso de Voz sobre IP.

8.6 Enfileiramento Custom Queueing

O algoritmo da fila CQ (Custom Queueing) permite especificar uma percentagem da banda para uma determinada aplicação (alocação absoluta da banda). A banda reservada é compartilhada proporcionalmente, no percentual pré-definido, entre as aplicações e os usuários. O restante da banda é compartilhado entre os outros tipos de tráfego.

O algoritmo CQ controla o tráfego alocando uma determinada parte da fila para cada fluxo classificado. As filas são ordenadas ciclicamente num esquema round-robin, onde, para cada fila, é enviado a quantidade de pacotes referente à parte da banda alocada antes de passar para a fila seguinte. Associado a cada fila, há um contador configurável que estabelece quantos bytes devem ser enviados antes da passar para a próxima fila.

Até 17 filas podem ser definidas, mas a fila zero é reservada para mensagens do sistemas como sinalização, keep-alive, etc. A classificação CQ pode ser feita por endereço fonte ou destino, por protocolo (IP, IPX, Appletalk, SNA, DecNet, etc), por precedência IP, por interface de entrada e ainda por listas de acesso.

8.7 Controle de congestionamento em sub-redes de circuito virtual (Controle Dinâmico)

Uma técnica que é amplamente utilizada para impedir que um congestionamento que já tenha começado piore é o controle de admissão (admission control). A idéia é simples: Uma vez que o congestionamento tenha sido detectado, nenhum outro circuito virtual será estabelecido até a resolução do problema. Logo, todas as tentativas de estabelecer novas conexões da camada de transporte falharão.

Uma estratégia alternativa é permitir novos circuitos virtuais, mas rotear com cuidado todos os novos circuitos virtuais em áreas problemáticas.

Uma outra estratégia relacionada a circuitos virtuais é negociar um acordo entre o host e a sub-rede quando um circuito virtual for estabelecido. Sendo assim a sub-rede reserva recursos. Uma desvantagem de fazer isso, o tempo todo é a tendência ao desperdício de recursos.

8.8 Pacote Regulador

Essa estratégia pode ser utilizada em sub-redes de circuito virtual ou de datagramas. Cada roteador pode monitorar facilmente a utilização de suas linhas de saída e de outros recursos. Cada pacote recém-chegado é conferido para sabermos se sua linha de entrada encontra-se em estado de advertência. Se estiver, o roteador enviará um pacote regulador (choke packet) para o host de origem, informando o destino encontrado no pacote. O pacote original é marcado (um bit de cabeçalho é ativado) para que não gere mais pacotes reguladores ao longo do caminho e, depois será encaminhado da forma usual.

Quando o host de origem obtém o pacote regulador, é necessário reduzir o tráfego enviado para o destino especificado em X por cento. Como outros pacotes com o mesmo destino já estão provavelmente em andamento vão gerar ainda mais pacotes reguladores, o host deverá ignorar os pacotes reguladores que se refiram a esse destino por um intervalo de tempo fixo. Depois que esse tempo tiver expirado, o host detectará mais pacotes reguladores para outro intervalo. Se chegar um, a linha ainda estará congestionada; logo, o host reduzirá o fluxo ainda mais e começará a ignorar pacotes reguladores novamente. Se não chegar pacote regulador algum durante o período de detecção, o host poderá aumentar o fluxo novamente. O feedback implícito desse protocolo pode ajudar a evitar congestionamento sem estrangular o fluxo, a menos que ocorra algum problema.

Um problema com o uso de pacotes regulares é que a ação a ser realizada pelos hosts de origem é voluntária, logo, suponha que um roteador esteja sendo atolado por pacotes de quatro origens e envie pacotes reguladores a todas elas. Se somente uma delas interromper a transmissão, o resultado é que o host honesto obtém uma parte menor da largura da banda que tinha antes.

Para contornar este problema, Nagle (1987) propôs um algoritmo de enfileiramento ponderado a seguir.

8.9 Enfileiramento com justiça ponderado (Weighted Fair queuing).

A essência do algoritmo é que os roteadores têm várias filas para cada linha de saída, uma para cada origem. Quando uma linha se torna inativa o roteador percorre as filas alternadamente, pegando o primeiro pacote da fila seguinte. Dessa forma, com n hosts disputando uma determinada linha de saída, cada host conseguirá enviar um a cada n pacotes.

Um problema com esse algoritmo é que ele dá a todos os hosts a mesma prioridade. Logo modificou-se o algoritmo para que ele possa dar prioridades diferentes aos servidores de arquivos e a outros servidores aumentando o tamanho da banda de modo que eles possam receber dois ou mais bytes por pulso.

Existe também os **pacotes reguladores hop a hop** para as redes de altas velocidades e longas distâncias. Neste método o pacote regulador tem efeito a cada hop por onde ele passa.

O efeito desse esquema hop a hop é oferecer alívio rápido no ponto de congestionamento ao preço do consumo de mais buffer do fluxo ascendente. Dessa maneira, o congestionamento pode ser cortado pela raiz sem que nenhum pacote seja perdido.

8.10 Escoamento de Carga

Escoamento de carga é uma maneira diferente de dizer que, quando os roteadores estão sendo inundados por pacotes que não podem manipular, eles simplesmente os jogam fora. Um roteador que está se afogando com pacotes pode simplesmente selecionar aleatoriamente aqueles que deverão ser descartados, mas normalmente é possível algo melhor que isso. O pacote a ser descartado pode depender das aplicações em execução. Descartar os pacotes com inteligência requer cooperação dos transmissores. Para implementar uma política de descarte inteligente, as aplicações devem marcar seus pacotes em classe de prioridade para indicar a importância. Se fizerem isso, quando chegar o momento de descartar os pacotes, os roteadores poderão primeiro descartar os pacotes da classe mais baixa e depois os da classe mais baixa seguinte e assim sucessivamente.

8.11 Controle de Congestionamento por Multicast

RSVP - Resource reSerVation Protocol

Esse protocolo permite que diversos transmissores transmitam para vários grupos de receptores, possibilita que receptores individuais mudem livremente de canal e otimiza o uso da largura da banda enquanto elimina o congestionamento.

Em seu formato mais simples, o protocolo utiliza roteamento multicast com spanning trees. Cada grupo recebe um endereço de grupo. Para enviar dados a um grupo, um transmissor coloca o endereço desse grupo nos pacotes enviados. Em seguida, o algoritmo de roteamento multicast padrão constrói uma spanning tree que cobre todos os membros. O algoritmo de roteamento não faz parte do RSVP. A única diferença em relação ao multicast normal (gerenciar fluxos de diversas origens para vários destinos) são algumas informações extras para que o multicast de grupo periodicamente forneça para os roteadores ao longo da árvore, a fim de manter estruturas de dados em suas respectivas memórias.

Ao fazer uma reserva, um receptor pode (opcionalmente) especificar uma ou mais origens a partir das quais deseja receber informações. Ele também pode especificar se essas opções serão fixas durante o período de reserva, ou se o receptor deseja manter em aberto a opção de alterar as origens, x para usá-la mais tarde. Os roteadores utilizam mais informações para otimizar o planejamento da largura de banda. Especificamente, dois receptores só são

configurados para compartilhar um caminho se ambos concordarem em não alterar as origens posteriormente.

O motivo para esta estratégia no caso totalmente dinâmico é que a largura de banda reservada é desacoplada da opção de origem. Quando reserva largura de banda, um receptor pode alternar para outra origem e manter parte do caminho existente que seja válida para a nova origem. Exemplo, se o host 2 estiver transmitindo diversos fluxos de vídeo, por exemplo, o host 3 poderá alternar entre eles, quando quiser, sem alterar sua reserva, pois os roteadores não se importam com o programa a que o receptor está assistindo.

9. RTP- Real-Time Transport Protocol

O RTP é o protocolo padrão para o transporte de fluxos de tempo-real, como áudio e vídeo. Ele é constituído de uma parte de dados e uma parte de controle (RTCP).

A parte de dados do RTP é um protocolo que garante suporte a aplicações com propriedades de tempo-real, realizando a reconstrução de temporizações, a detecção de perdas e identificação de conteúdo. O RTCP provê suporte à conferência de grupos de qualquer tamanho na Internet. Este suporte inclui a identificação de fontes e suporte a roteadores de nível 2 (*bridges*) para áudio e vídeo, bem como tradutores de multicast para unicast. É oferecida uma realimentação de QoS dos receptores para o grupo multicast, além do suporte à sincronização de fluxos de mídias diferentes.

Como forma de diminuir o sobrecarga de processamento gerada pelos cabeçalhos IP-UDP-RTP, que pode chegar a 40 octetos, é adotada versão CRTP, com compressão do cabeçalho para um valor de 2 a 4 octetos.

9.1 Compressão RTP

O protocolo CRTP (Compressed Real-time Transport Protocol) comprime o cabeçalho do pacote RTP, que transporta o tráfego de voz. O campo de dados (payload) de um pacote IP de voz é composto pelo pacote RTP, que transporta o sinal de voz propriamente dito, encapsulado num pacote UDP.

O tamanho do cabeçalho RTP é de 12 bytes, o do UDP de 8 bytes e do IP de 20 bytes, totalizando 40 bytes. Considerando o payload IP de 20 bytes, só de cabeçalho teríamos 66,66% da banda da conexão.

Aí é que entra o protocolo CRTP, que, na maior parte do tempo, consegue uma compressão de cabeçalho de 40 para 2 bytes, ou para 4 se considerarmos o checksum UDP. Essa compressão corresponde a uma redução de até 95% na sobrecarga (overhead) referente aos cabeçalhos.

O tráfego total destinado a uma determinada interface é classificado, e o que for RTP é separado para compressão. O tráfego RTP é, então, processado num compressor e colocado novamente na fila para ser transmitido.

Nos gateways VoIP Cisco, essa compressão é habilitada na interface serial, como mostrado no exemplo abaixo:

```
interface serial2/1
```

```
    ip rtp header-compression
```

```
    encapsulation ppp
```

```
ip rtp compression-connections 25
```

A compressão CRTP é bastante útil em conexões de baixa velocidade (64 Kbps, p. ex.), mas também possui desvantagens. Se for habilitada compressão RTP, o roteador só funcionará em process switch em vez de fast switch, o que, combinado com a sobrecarga da execução CRTP, pode causar queda de desempenho do mesmo. Portanto, o uso do CRTP não é sugerido em enlaces de alta velocidade, uma vez que a relação custo versus benefício pode não compensar.

Biografia:

- Livro: Computer Networks - Third Edition
Autor: Andrew S. Tanenbaum
- Apostila de Redes de Alta Velocidade - Apostila 1
Autores: Eleri Cardozo e Mauricio F. Magalhães
2000 - 2001 - 1 de Agosto de 2001
- Artigo: Priorização de Tráfego em redes IP
Autor: Fábio Ribas Ardeola
- Artigo: Qualidade de Serviço (VoIP) - Parte 1
Autor: Adailton J. S. Silva
12/05/2000
RNP - NewsGeneration