Entender a QoS nos Switches da família Catalyst 6000

Contents

Introduction

Este documento explica as características de qualidade de serviço (QoS) disponíveis nos switches da família Catalyst 6000. Este documento abrange os recursos de configuração da qualidade de serviço (QoS) e fornece alguns exemplos de como a QoS pode ser implementada.

Este documento não é um guia de configuração. Exemplos de configuração são usados neste documento para auxiliar na explicação dos recursos de QoS do hardware e software da família Catalyst 6000. Para referência de sintaxe para estruturas de comando de QoS, consulte os seguintes guias de configuração e de comando para a família Catalyst 6000:

- Switches da família Catalyst 6500

Definindo a QoS da camada 2

Ainda que muitos possam pensar que o QoS de Switches na camada 2 (L2) esteja relacionado apenas à priorização das estruturas de Ethernet, alguns percebem que isto envolve muito mais coisas. A QoS L2 inclui o seguinte:

1. **Programação da fila de entrada:** quando o quadro entra na porta, pode ser atribuído a um de um número de filas com base em portas antes de ser programado para Switching em uma porta de saída. Geralmente, várias filas são usadas onde tráfego diferente exige níveis de serviço diferentes ou onde a latência do switch deve ser mantida ao

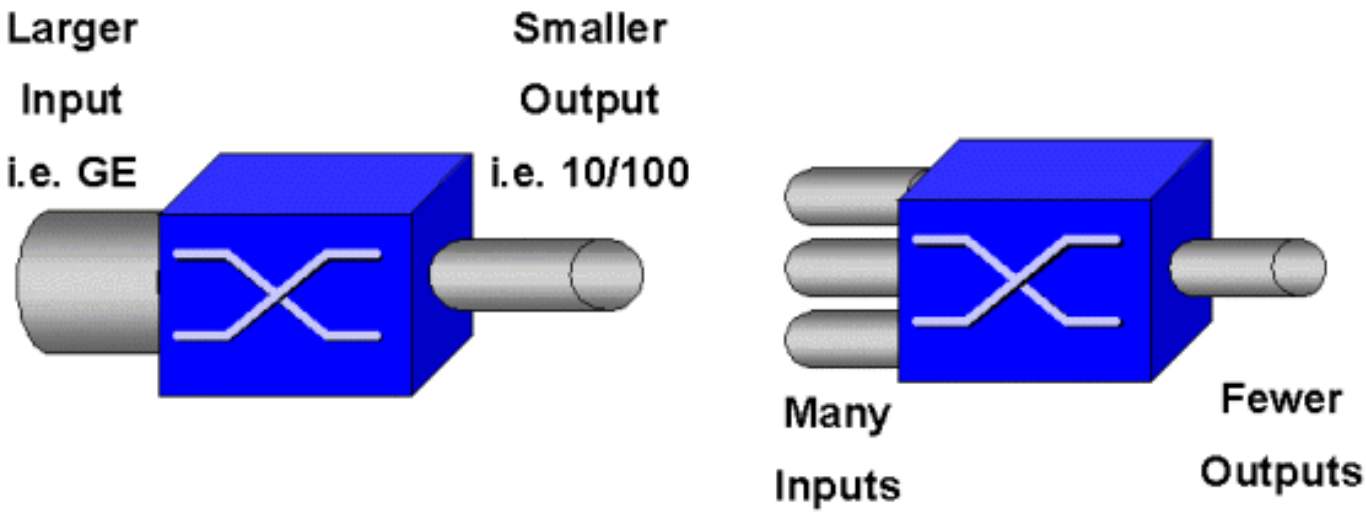
mínimo. Por exemplo, dados de voz e vídeo baseados em IP exigem baixa latência, portanto, pode haver necessidade de trocar esses dados antes de comutar outros dados, como FTP, Web, e-mail, Telnet e assim por diante.

2. **Classificação** o processo de classificação envolve a inspeção de diferentes campos no cabeçalho Ethernet L2, juntamente com campos no cabeçalho IP (Camada 3 (L3)) e no cabeçalho Transmission Control Protocol/User Datagram Protocol (TCP/UDP) (Camada 4 (L4))) para ajudar a determinar o nível de serviço que será aplicado ao quadro à medida que ele transita pelo switch.
3. **Vigilância:** policiamento é o processo de inspeção de um quadro Ethernet para ver se ele excedeu uma taxa de tráfego predefinida em um determinado período (normalmente, esse período é um número fixo interno do switch). Se esse quadro estiver fora de perfil (ou seja, faz parte de um fluxo de dados além do limite de taxa predefinido), ele pode ser descartado ou o valor de Classe de Serviço (CoS) pode ser marcado.
4. **Regravando:** O processo de reescrever é a capacidade do Switch de modificar CoS no cabeçalho Ethernet de bits ToS (Type of Service) no cabeçalho IPV4.
5. **Programação de fila de saída:** após os processos de regravação, o Switch colocará o quadro Ethernet em uma fila de saída apropriada (de saída) para switching. O Switch executará o gerenciamento de buffer nesta consulta, assegurando que o buffer não crie sobrefluxos. Geralmente, ele fará isso utilizando um algoritmo de descarte antecipado aleatório (RED - Random Early Discard), em que os quadros aleatórios são removidos (descartados) da fila. O RED ponderado (WRED) é um derivado do RED (usado por determinados módulos da família Catalyst 6000), pelo qual os valores de CoS são inspecionados para determinar quais quadros serão descartados. Quando os buffers atingem limites predefinidos, quadros com prioridade menor são normalmente descartados, mantendo os quadros com a prioridade maior na fila.

Este documento explica com mais detalhes cada um dos mecanismos acima e como eles se relacionam com a família Catalyst 6000 nas seções a seguir.

A necessidade de QoS em um Switch

Backplanes enormes, milhões de pacotes comutados por segundo e switches sem bloqueio são todos sinônimos de muitos switches atualmente. Por que um QoS é necessário? A resposta é devido a um congestionamento.



Um switch pode ser o switch mais rápido do mundo, mas se você tiver um dos dois cenários mostrados na figura acima, esse switch terá congestionamento. Em momentos de congestionamento, se os recursos de gerenciamento de congestionamento não estiverem instalados, os pacotes serão descartados. Quando os pacotes são descartados, as retransmissões ocorrem. Ao ocorrerem retransmissões, a carga da rede pode aumentar. Em redes já congestionadas, isso pode aumentar os problemas de desempenho existentes e, potencialmente, diminuir ainda mais o desempenho.

Com redes convergentes, o gerenciamento do congestionamento é ainda mais crítico. O tráfego sensível à latência, como voz e vídeo, pode ser seriamente afetado se ocorrerem atrasos. A simples adição de mais buffers a um switch também não necessariamente aliviará problemas de congestionamento. O tráfego sensível à latência precisa ser comutado o mais rápido possível. Primeiro, você precisa identificar esse tráfego importante por meio de técnicas de classificação e depois implementar técnicas de gerenciamento de buffer para evitar que o tráfego de prioridade mais alta seja descartado durante o congestionamento. Finalmente, é necessário incorporar técnicas de agendamento para comutar pacotes importantes das filas o mais rápido possível. Como você lerá neste documento, a família Catalyst 6000 implementa todas essas técnicas, tornando seu subsistema de QoS uma das mais abrangentes do setor atualmente.

Todas as técnicas de QoS descritas na seção anterior serão exploradas com mais detalhes neste documento.

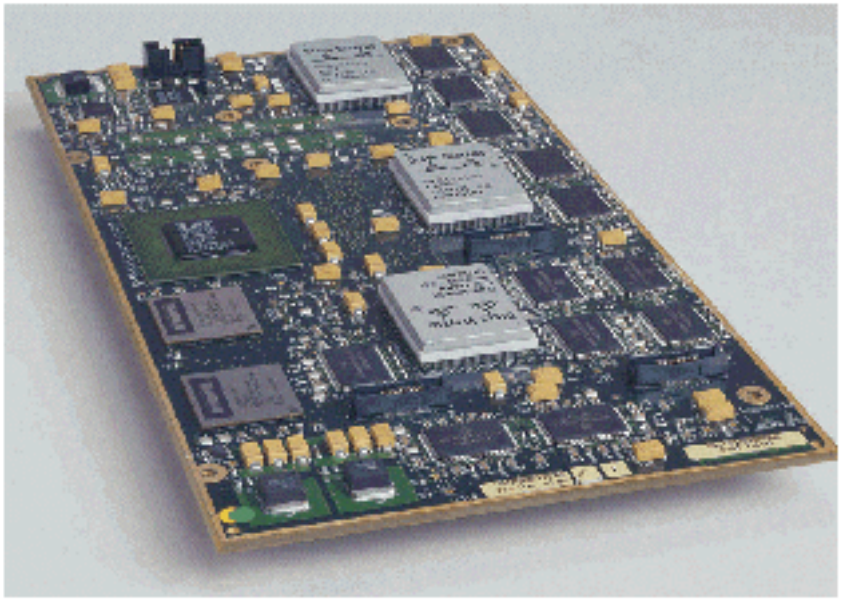
Suporte de hardware para QoS no Catalyst 6000 Family

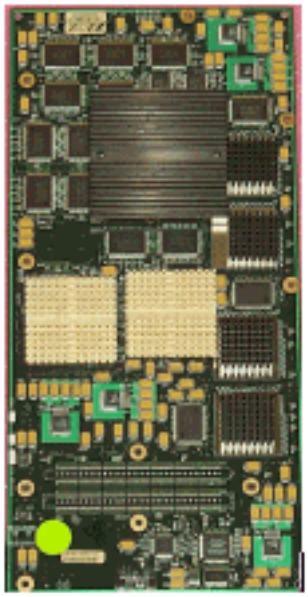
Para suportar QoS na família Catalyst 6000, é necessário algum suporte de hardware. O hardware que suporta QoS inclui o Multilayer Switch Feature Card (MSFC), o Policy Feature Card (PFC) e os Port Application Specific Integrated Circuits (ASICs) nas próprias placas de linha. Este documento não abordará as capacidades de QoS do MSFC; em vez diz, se concentrará nas capacidades de QoS do PFC e nos ASICs das placas de linha.

PFC

O PFC versão 1 é uma placa secundária acomodada no Supervisor I (SupI) e o Supervisor IA (SupIA) da família Catalyst 6000. O PFC2 é versão aprimorada do PFC1 e é fornecido com o novo Supervisor II (SupII) e alguns ASICs novos integrados na placa. Embora o PFC1 e o PFC2 sejam conhecidos principalmente por sua aceleração de hardware do switching L3, o QoS é um

de seus outros propósitos. Os PFCs são mostrados abaixo.





Embora PFC 1 e PFC2 sejam essencialmente iguais, existem algumas diferenças na funcionalidade QoS. Ou seja, o PFC2 adiciona o seguinte:

1. A capacidade de aplicar a política de QoS em uma DFC (Placa de encaminhamento distribuído).

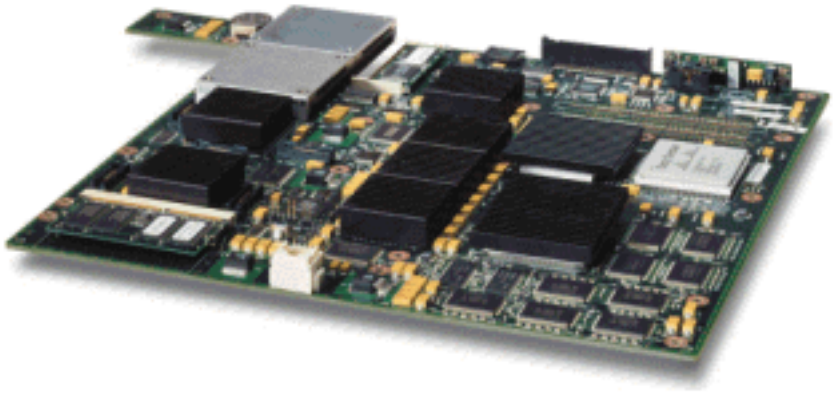
2. As decisões sobre vigilância são ligeiramente diferentes. O PFC1 e o PFC2 suportam vigilância normal em que os quadros são descartados ou marcados como inativos se uma política de agregação ou microfluxo retorna uma decisão fora de perfil. No entanto, o PFC2 adiciona suporte para uma taxa de excesso, o que indica um segundo nível de vigilância no qual as ações de política podem ser tomadas.

Quando um vigilante de taxa de excesso é definido, os pacotes podem ser descartados ou marcados para baixo quando excedem a taxa de excesso. Se um nível de polícia em excesso for definido, o mapeamento de DSCP em excesso será usado para substituir o valor de DSCP original por um valor marcado para baixo. Se apenas um nível de polícia normal for definido, o mapeamento DSCP normal será usado. O nível policial em excesso terá precedência para a seleção de regras de mapeamento quando ambos os níveis forem definidos.

Éimportante observar que as funções de QoS descritas neste documento executadas pelos ASICs mencionados geram altos níveis de desempenho. O desempenho de QoS em um Catalyst 6000 Family básica (sem módulo de Switch Fabric) produz 15 MPPS. Ganhos adicionais de desempenho podem ser obtidos para QoS se os DFCs forem usados.

DFC

O DFC pode ser anexado ao WS-X6516-GBIC como uma opção. No entanto, é uma instalação padrão na placa WS-X6816-GBIC. Ele também pode ser suportado em outras placas de linha de malha futuras, como a placa de linha de malha 10/100 (WS-X6548-RJ45), placa de linha de malha RJ21 (WS-X6548-RJ21) e a placa de linha 100FX (WS-X6524) MM-FX). O DFC é mostrado a seguir.



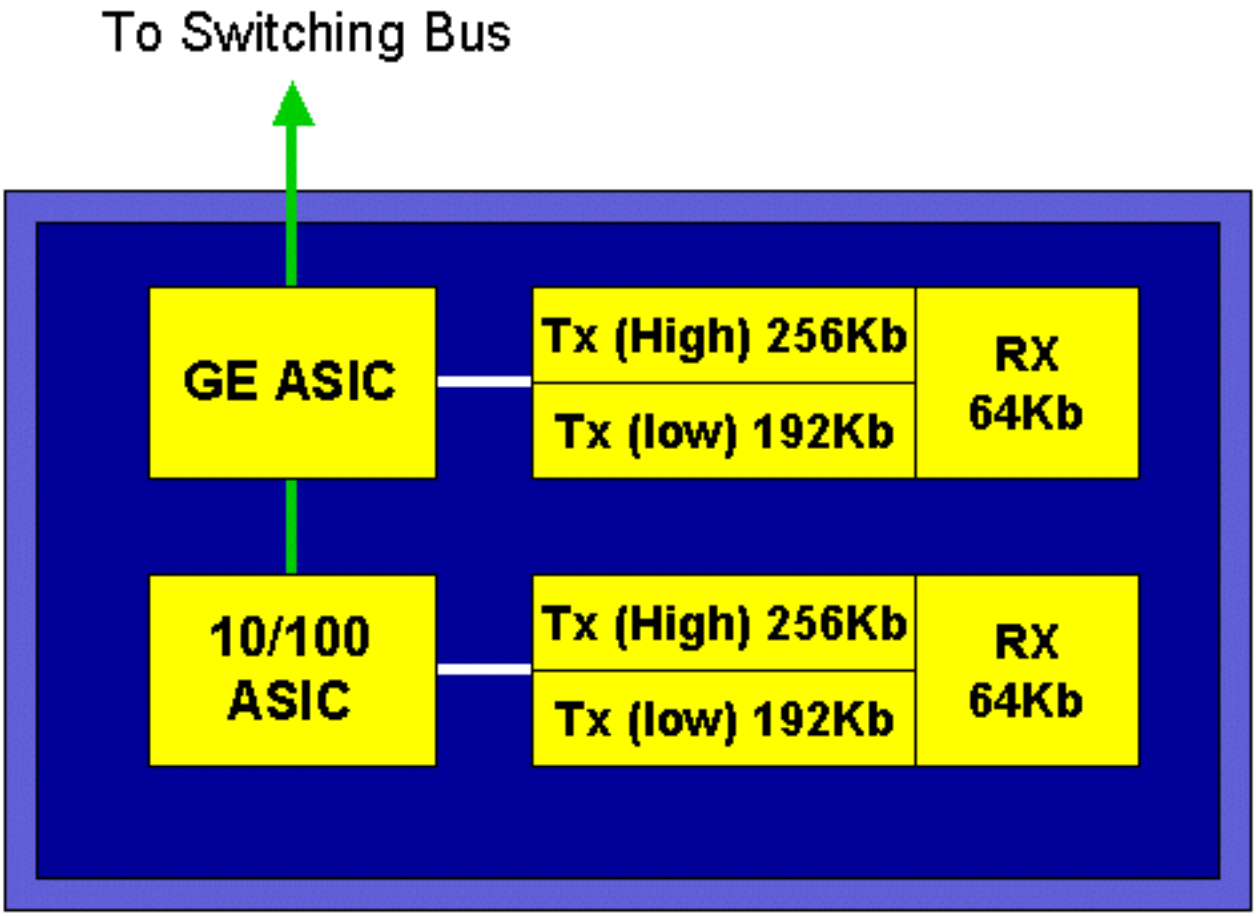
A DFC permite que a placa de linha de estrutura (conexão de barra cruzada) execute o switching local. Para fazer isso, ele também deve suportar qualquer política de QoS que tenha sido definida para o switch. O administrador não pode configurar diretamente o DFC; em vez disso, ele está sob o controle do MSFC/PFC mestre no supervisor ativo. O PFC principal empurrará uma tabela FIB (Forwarding Information Base), que fornece ao DFC suas tabelas de encaminhamento L2 e L3. Também será enviada uma cópia das políticas de QoS para que elas também sejam locais para a placa de linha. Subsequentemente, as decisões de comutação local podem fazer referência à cópia local de qualquer política de QoS que forneça velocidades de processamento de QoS de hardware e forneça níveis mais altos de desempenho através de comutação distribuída.

ASICs baseados em portas

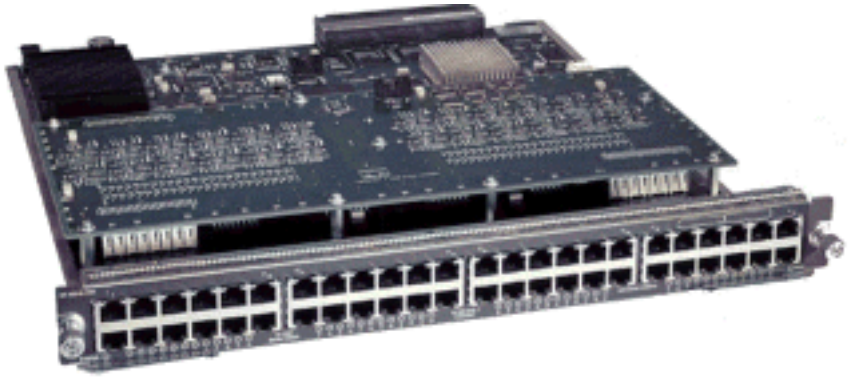
Para completar a imagem do hardware, cada uma das placas de linha implementam um número de ASICs. Esse ASICs implementam as filas, a colocação em buffer e os limiares utilizados para o armazenamento temporário de quadros enquanto atravessam o Switch. Nas placas 10/100, uma combinação de ASICs é usada para suprir as 48 portas 10/100.

Placas de linha 10/100 originais (WS-X6348-RJ45)

OS ASICs de 10/100 fornecem uma série de filas de Recepção (Rx) e Transmissão (Tx) para cada porta 10/100. Os ASICs fornecem 128 K de buffer por porta 10/100. Consulte as notas de versão para obter detalhes sobre o que está disponível em cada buffer de porta em cada placa de linha. Cada porta nesta placa de linha suporta uma fila Rx e duas filas TX identificadas como alta e baixa. Isso é mostrado no diagrama abaixo.



No diagrama acima, cada ASIC 10/100 fornece uma divisão para 12 portas 10/100. Para cada porta 10/100, são fornecidos buffers de 128 K. Os 128 K de buffers são divididos entre cada uma das três filas. As figuras mostradas na fila acima não são o padrão, entretanto elas podem ser uma representação do que poderia ser configurado. A fila Rx única fica com 16 K e a memória restante (112 K) é dividida entre as duas filas Tx. Por padrão (no CatOS), a fila alta ganha 20% desse espaço e a fila baixa recebe 80%. No Catalyst IOS, o padrão é dar à fila alta 10% e à fila baixa 90%.

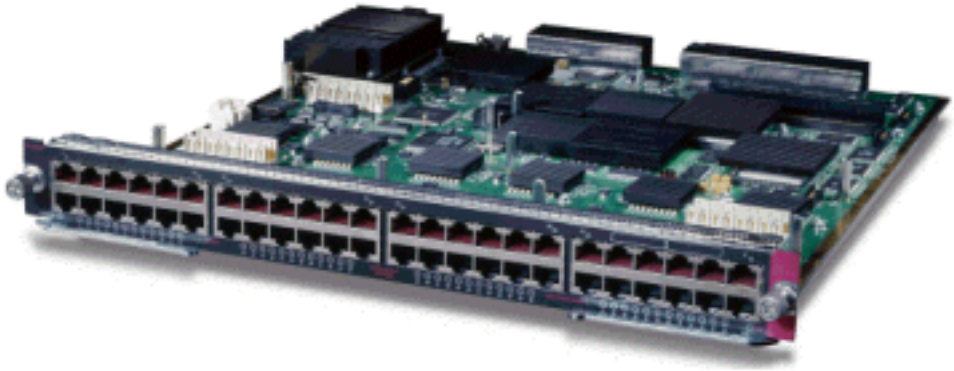


Embora a placa ofereça buffer de estágio duplo, somente o buffer baseado em ASIC 10/100 está disponível para ser manipulado durante a configuração de QoS.

Placas de linha Fabric 10/100 (WS-X6548-RJ45)

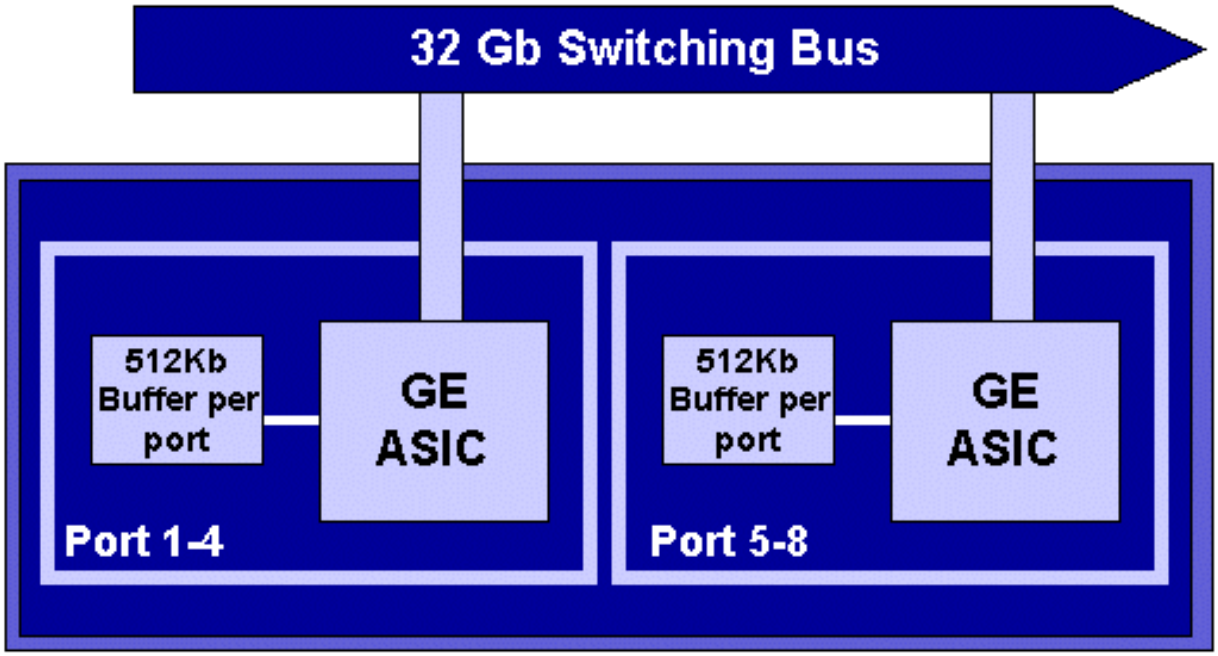
As novas ASICs de 10/100 fornecem uma série de filas Rx e TX para cada porta 10/100. Os ASICs fornecem um pool compartilhado de memória disponível nas portas 10/100. Consulte as notas de versão para obter detalhes sobre o que está disponível em cada buffer de porta em cada placa de linha. Cada porta nesta placa de linha suporta duas filas Rx e três filas TX. Cada fila Rx

e uma fila TX são identificadas como uma fila de prioridade absoluta. Ela age como uma fila de latência baixa, que é ideal para tráfego sensível à latência, como o tráfego de Voz sobre IP (VoIP).

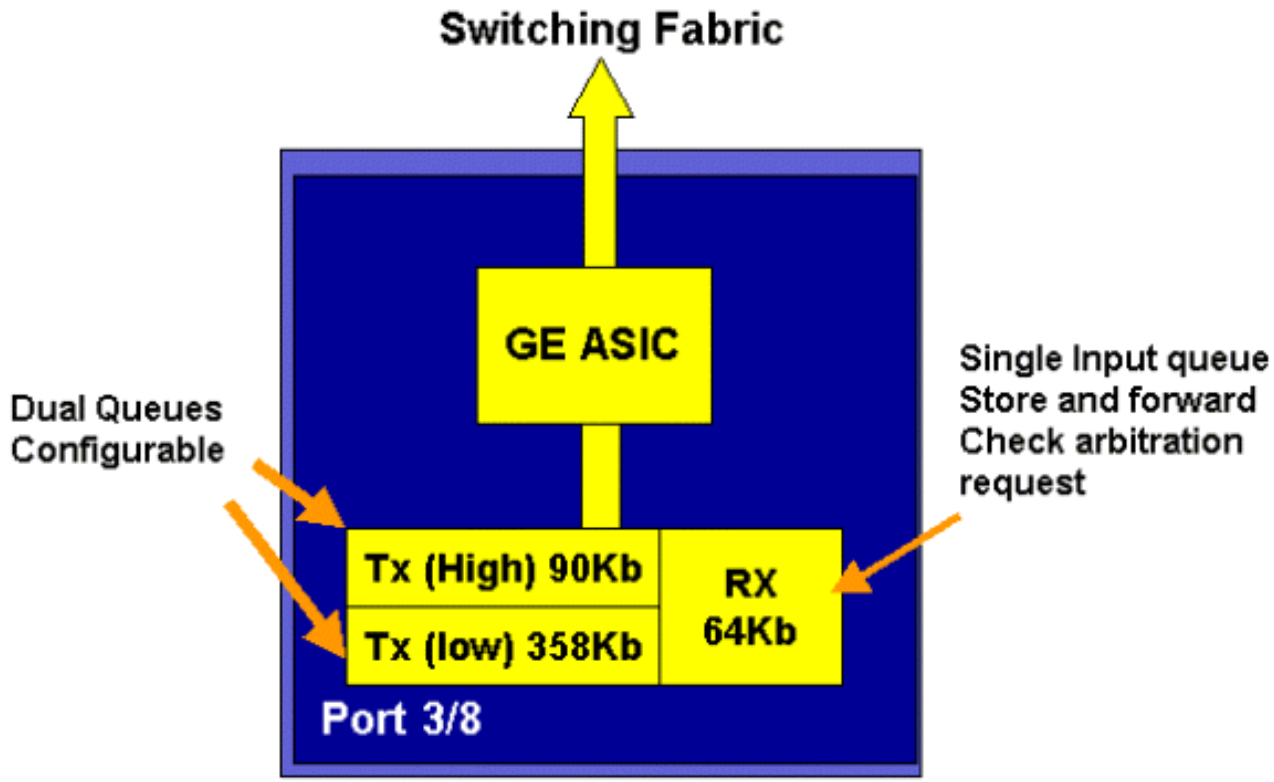


Placas de linha GE (WS-X6408A, WS-X6516, WS-X6816)

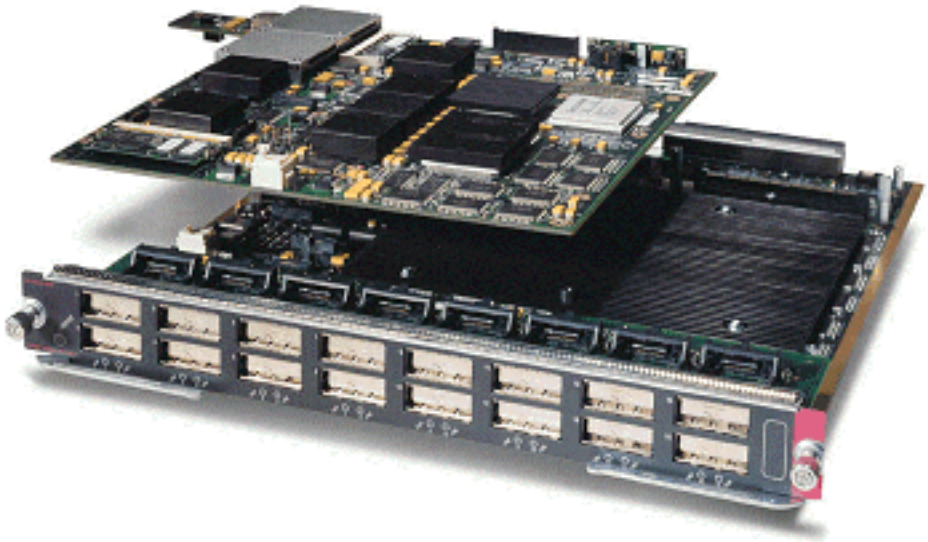
Para placas de linha GE, o ASIC fornece 512 K de armazenamento em buffer por porta. Uma representação da placa de linha GE de oito portas é mostrada no diagrama abaixo.



Assim como nas portas 10/100, cada porta GE tem três filas, uma Rx e duas filas TX. Este é o padrão da placa de linha WS-X6408-GBIC, e está mostrado no diagrama abaixo.



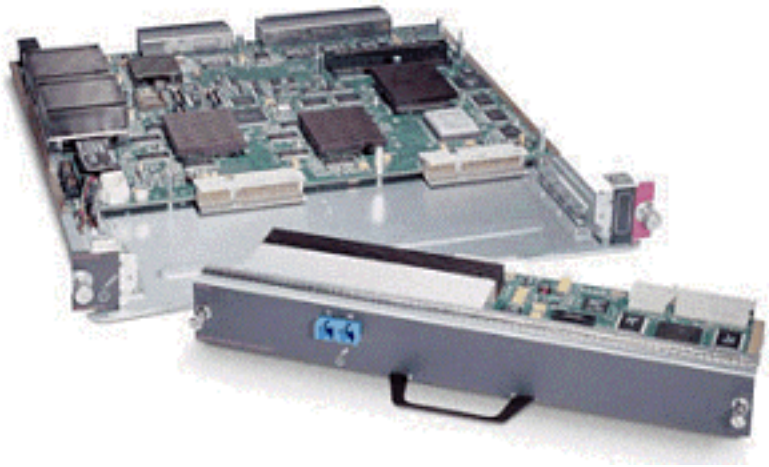
Nas placas GE de 16 portas mais recentes, nas portas GBIC em SupIA e SupII e na placa GE de 8 portas WS-X6408A-GBIC, são fornecidas duas filas extras de prioridade máxima (SP). Uma das filas SP é atribuída como uma fila Rx e a outra, como uma fila TX. Essa fila SP é usada principalmente para enfileirar tráfego sensível à latência, como voz. Com a fila SP, qualquer dado colocado nessa fila será processado antes do dado nas filhas alta e baixas. Somente quando a fila do SP estiver vazia as filas alta e baixa serão atendidas.



Placas de linha 10 GE (WS-X6502-10GE)

No último semestre de 2001, a Cisco apresentou um conjunto de placas de linha 10 GE fornecendo uma porta de 10 GE por placa de linha. Este módulo usa um slot do chassi 6000. A placa de linha 10 GE suporta QoS. Para a porta 10 GE, ela fornece duas filas Rx e três filas TX. Cada fila Rx e uma fila TX são designadas como uma fila SP. O buffer também é fornecido para a porta, fornecendo um total de 256 K de buffer Rx e 64 MB de buffer TX. Esta porta implementa

uma estrutura de fila de 1p1q8t para o lado Rx e uma estrutura de fila de 1p2q1t para o lado TX. As estruturas de fila são detalhadas adiante neste documento.



Resumo de hardware de QoS da família Catalyst 6000

Os componentes de hardware que executam as funções de QoS acima na família Catalyst 6000 estão detalhados na tabela abaixo.

QoS Process	Catalyst 6500 Component that performs function
Input Scheduling	Performed by port ASIC's L2 only with or without the PFC
Classification	Performed by Supervisor or PFC L2 only is done by Supervisor L2/3 is done by PFC
Policing	Done by PFC via L3 forwarding Engine
Packet Re-write	Done by port ASIC's L2/L3 based on classification done in point 2 above
Output Scheduling	Done by port ASIC's L2/L3 based on classification done in point 2 above

Suporte para QoS ao software da família Catalyst 6000

A família Catalyst 6000 suporta dois sistemas operacionais. A plataforma do software original, CatOS, derivou-se da base de código usada na plataforma do Catalyst 5000. Mais recentemente, a Cisco apresentou o Integrated Cisco IOS ® (modo nativo) (anteriormente conhecido como Native IOS), que usa uma base de código derivada do Cisco Router IOS. Ambas as plataformas de SO (CatOS e Cisco IOS Integrado (Modo Nativo)) implementam o suporte de software para ativar a QoS na plataforma da família de switches Catalyst 6000 usando o hardware descrito nas seções anteriores.

Note: Este documento utiliza exemplos de configuração de ambas as plataformas de SO.

Mecanismos de prioridade em IP e Ethernet

Para que qualquer serviço de QoS seja aplicado aos dados, deve haver uma maneira de marcar ou priorizar um pacote IP ou um quadro Ethernet. Os campos ToS e CoS são usados para isso.

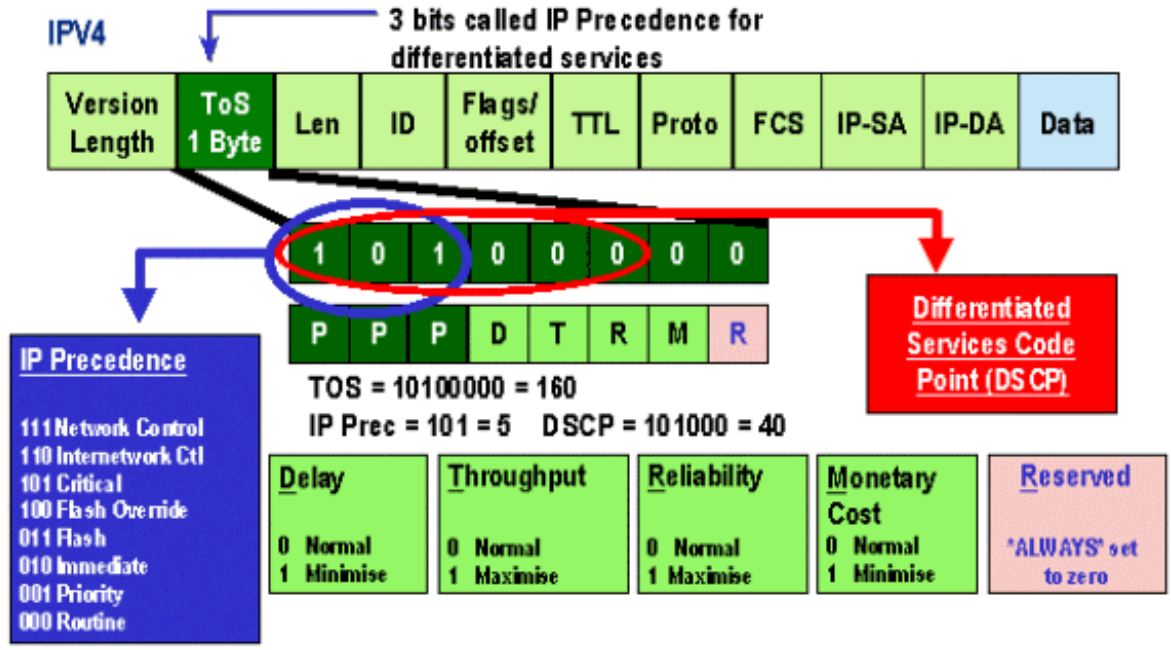
ToS

ToS é um campo de um byte que existe em um cabeçalho IPV4. O campo ToS consiste em oito bits, dos quais os três primeiros são utilizados para indicar a prioridade do pacote IP. Esses três primeiros bits são mencionados como os bits de precedência de IP. Esses bits podem ser definidos de zero a sete, sendo zero a prioridade mais baixa e sete a prioridade mais alta. O suporte está disponível para definir a presença IP no IOS por vários anos. O suporte a reinicialização de precedência do IP pode ser proporcionado pelo MSFC ou pelo PFC (independente do MSFC). Uma configuração confiável de não confiável também pode apagar qualquer configuração de precedência de IP em um quadro de entrada.

Os valores que podem ser definidos para precedência de IP são os seguintes:

IP Precedence bits	IP Precedence Value
000	Routine
001	Priority
010	Intermediate
011	Flash
100	Flash Override
101	Critical
110	Internetwork Control
111	Network Control

O diagrama abaixo representa os bits de precedência do IP no cabeçalho de ToS. Os três Bits mais significativos (MSB) são interpretados como bits de precedência de IP.



Mais recentemente, o uso do campo ToS foi ampliado para abranger os seis MSBs, conhecidos como DSCP. O DSCP resulta em 64 valores de prioridade (dois à potência de seis) que podem

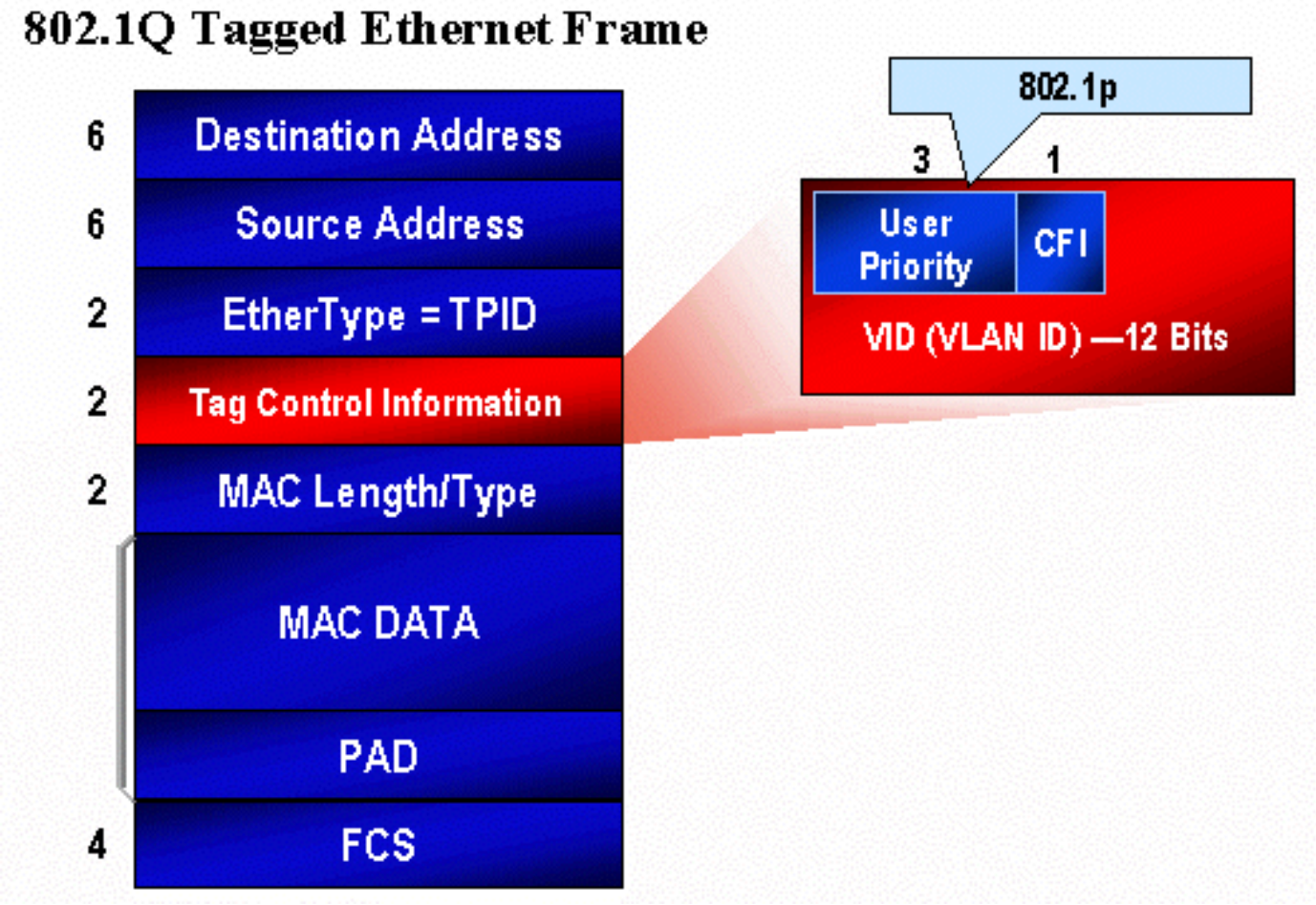
ser atribuídos ao pacote IP.

A família Catalyst 6000 pode manipular o ToS. Isso pode ser alcançado por meio do PFC e do MSFC. Quando um quadro aparece no Switch, um valor de DSCP é atribuído a ele. Este valor de SHCP é utilizado internamente no Switch para atribuir níveis de serviço (políticas de QoS) definidos pelo administrador. O DSCP pode já existir em um quadro e ser utilizado ou o DSCP pode ser derivado de um CoS existente, precedência IP ou DSCP no quadro (a porta deve ser confiável). Um mapa é usado internamente no switch para derivar o DSCP. Com oito valores possíveis de precedência de CoS/IP e 64 valores possíveis de DSCP, o mapa padrão irá mapear CoS/IPPrec 0 para DSCP 0, CoS/IPPrec 1 para DSCP 7, CoS/IPPrec 2 para DSCP 15 e assim por diante. Esses mapeamentos padrão podem ser substituídos pelo administrador. Quando o quadro estiver programado para uma porta de saída, o CoS pode ser regravado e o valor de DSCP é usado para derivar o novo CoS.

CoS

CoS refere-se a três bits em um cabeçalho ISL ou em um cabeçalho 802.1Q que são usados para indicar a prioridade do quadro Ethernet à medida que ele passa por uma rede comutada. Para os fins deste documento, falamos apenas do uso do cabeçalho 802.1Q. Os bits CoS do cabeçalho 802.1Q são comumente chamados de bits do 802.1p. Não surpreendentemente, há três bits de CoS, que correspondem ao número de bits usados para precedência de IP. Em muitas redes, para manter a QoS de ponta a ponta, um pacote pode atravessar os domínios L2 e L3. Para manter o QoC, ToS pode ser mapeado para CoS e vice-versa.

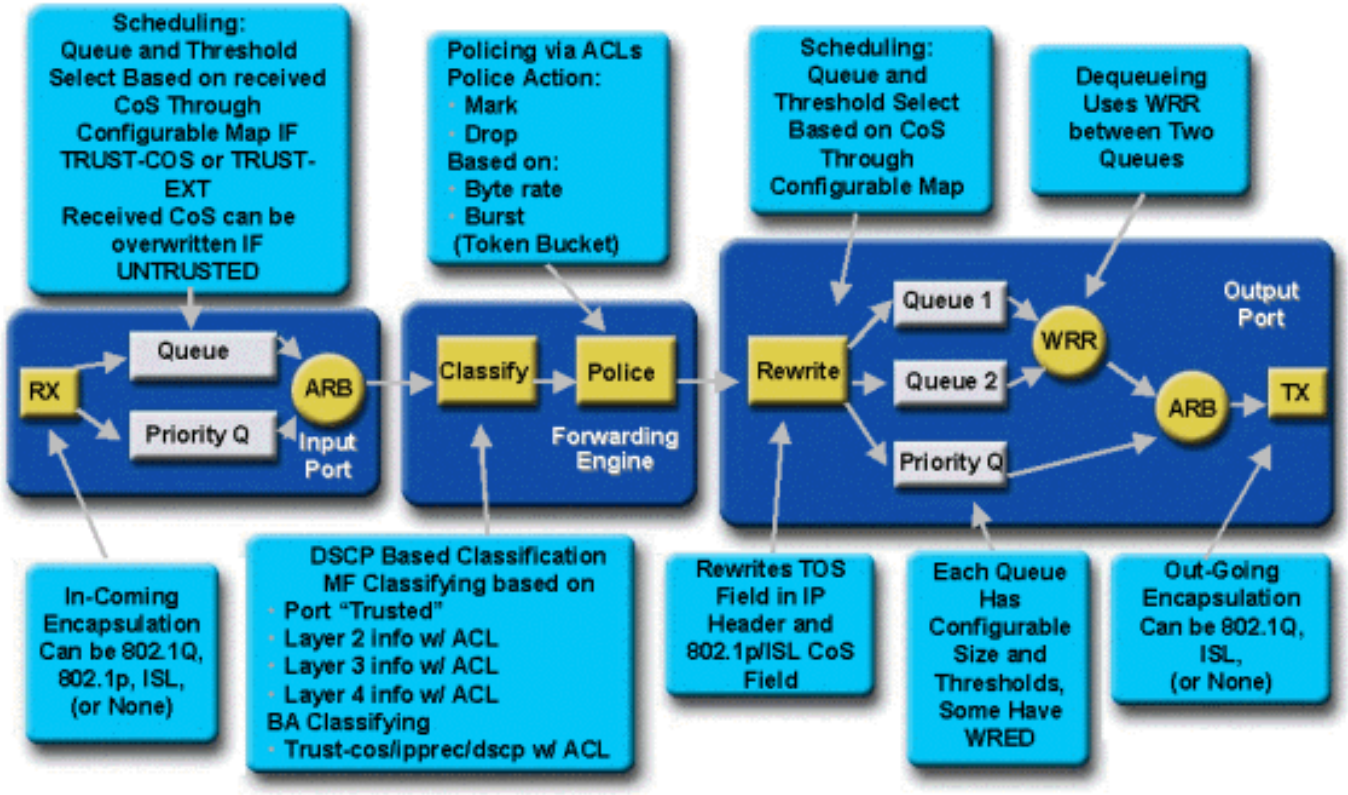
O diagrama abaixo é uma estrutura de Ethernet rotulada com um campo 802.1Q, que consiste em um Ethertipo de dois bytes e de um rótulo de dois bytes. Dentro da marca de dois bytes estão os bits de prioridade do usuário (conhecidos como 802.1p).



Fluxo de QoS no Catalyst 6000 Family

QoS no Catalyst 6000 Family é a implementação mais abrangente de QoS em todos os Cisco Catalyst Switches atuais. As seções a seguir descrevem como os vários processos de QoS são aplicados a um quadro à medida que ele transita pelo switch.

Anteriormente neste documento, observou-se que há vários elementos de QOS que muitos switches L2 e L3 podem oferecer. Esses elementos são: classificação, programação da fila de entrada, política, regravação e programação da fila de saída. A diferença em relação à família Catalyst 6000 é que esses elementos de QoS são aplicados por um mecanismo da L2 que tem um insight nos detalhes da L3 e da L4 e também informações de cabeçalho da L2. o diagrama a seguir resume como a família Catalyst 6000 implementa estes elementos.



Um quadro entra no switch e é inicialmente processado pela porta ASIC que recebeu o quadro. Ele colocará o quadro em uma fila Rx. Dependendo da placa de linha da família Catalyst 6000, haverá uma ou duas filas Rx.

A porta ASIC usará os bits CoS como indicador da fila que deverá receber o quadro (se várias filas de entrada existirem). Se a porta for classificada como não confiável, o ASIC da porta poderá substituir os bits de CoS existentes com base em um valor predefinido.

O quadro é então passado ao mecanismo de encaminhamento L2/L3 (PFC), que o classificará e opcionalmente o vigiará (limite de taxa). A classificação é o processo de atribuir ao quadro um valor de DSCP, que é usado internamente pelo switch para processar o quadro. O DSCP será derivado de um dos seguintes itens:

1. Um conjunto de valores DSCP existente antes do quadro entrar no switch
2. Bits de precedência do IP já definidos no cabeçalho IPV4. Como há 64 valores de DSCP e apenas oito valores de precedência de IP, o administrador configurará um

mapeamento que é usado pelo switch para derivar o DSCP. Os mapeamentos padrão estão em vigor caso o administrador não configure os mapas.
3. The received CoS bits already set prior to the frame entering the Switch. Assim como ocorre com a precedência IP, existe um máximo de oito valores CoS, sendo que cada um deve ser mapeado para um dos valores 64 DSCP. This map can be configured or the Switch can use the default map in place.
4. Defina para o quadro utilizando um valor padrão de DSCP, normalmente atribuído por uma entrada de Lista de controle de acesso (ACL).

Depois que um valor de DSCP é atribuído ao quadro, a vigilância (limitação de taxa) é aplicada, caso exista uma configuração de vigilância. A vigilância limitará o fluxo de dados através do PFC, descartando ou diminuindo o tráfego que estiver fora de perfil. Fora do perfil é um termo usado para indicar que o tráfego excedeu um limite definido pelo administrador como a quantidade de bits por segundo que a PFC enviará. O tráfego fora de perfil pode ser reduzido ou o valor de CoS pode ser marcado. No momento, PFC1 e PFC2 não oferecem suporte à vigilância de entrada (taxa limite). O suporte para policiamento de entrada e de saída estará disponível com a versão de um novo PFC.

O PFC passará o quadro para a porta de saída para processamento. Neste ponto, um processo de reescrita é invocado para modificar os valores de CoS no quadro e o valor ToS no cabeçalho IPV4. Isso é derivado do DSCP interno. O frame será colocado em uma fila de transmissão no valor CoS correspondente, pronto para a transmissão. Enquanto o quadro estiver na fila, a porta ASIC monitorará os buffers e implementará o WRED para evitar o excesso de buffers. Um algoritmo de programação WRR é então usado para programar e transmitir quadros da porta de saída

Cada uma das seções abaixo explorará esse fluxo com mais detalhes, fornecendo exemplos de configuração para cada uma das etapas descritas acima.

Filas, Buffers, Limiares e Mapeamentos

Antes que a configuração de QoS seja descrita em detalhes, certos termos devem ser explicados mais a fundo para garantir que você entenda totalmente os recursos de configuração de QoS do switch.

Filas

Cada porta no switch tem uma série de filas de entrada e saída usadas como áreas de armazenamento temporário para dados. As placas de linha da família Catalyst 6000 implementam números diferentes de filas para cada porta. Geralmente, as filas são implementadas no ASICs do hardware para cada porta. Nas placas de ingresso da família Catalyst 6000 de primeira geração, a configuração típica era uma fila de entrada e duas filas de saída. Nas placas de linha mais novas (10/100 e GE), o ASIC implementa um conjunto extra de duas filas (uma entrada e uma saída), resultando em duas filas de entrada e três filas de saída. Essas duas filas extras são filas SP especiais usadas para tráfego de latência sensível como VoIP. São atendidos de maneira SP. Ou seja, se um quadro chegar na fila SP, a programação de quadros nas filas mais baixas será interrompido para processar o quadro na fila SP. A programação dos pacotes de recomeço de fila(s) inferior(es) será feita somente quando a fila de SP estiver vazia.

Quando um quadro chegar a uma porta (para entrada ou saída) em horários de congestionamento, ele será colocado em uma fila. A decisão sobre em qual fila a estrutura será colocada geralmente é feita com base no valor de CoS no cabeçalho de Ethernet da estrutura recebida.

Na saída, um algoritmo de programação será empregado para esvaziar a fila de TX (saída). WRR é a técnica utilizada para se alcançar isso. Para cada fila, uma ponderação é usada para determinar a quantidade de dados que será esvaziada da fila antes de passar para a próxima fila. A pesagem atribuída pelo administrador é um número de 1 a 255 e isso é atribuído a cada fila TX.

Buffers

Cada fila recebe uma determinada quantidade de espaço de buffer para armazenar dados de trânsito. A memória é residente no ASIC de porta e é dividida e alocada por porta. Para cada porta GE, o GE ASIC atribui 512 K de espaço de buffer. Para portas 10/100, o ASIC de porta reserva 64 K ou 128 K (dependendo da placa de linha) de buffer por porta. Esse espaço de buffer é então dividido entre a fila Rx (de ingresso) e as filas TX (de saída).

Limiares

Um aspecto da transmissão de dados normal é que, se um pacote for descartado, ele acabará sendo retransmitido (fluxos de TCP). Em momentos de congestionamento, isso pode aumentar a carga na rede e, potencialmente, fazer com que os buffers sobrecarreguem ainda mais. Como forma de garantir que os buffers não sobrecarreguem, o switch da família Catalyst 6000 emprega várias técnicas para evitar que isso aconteça.

Os limites são níveis imaginários atribuídos pelo switch (ou pelo administrador) que definem os pontos de utilização nos quais o algoritmo de gerenciamento de congestionamento pode começar a descartar dados da fila. Nas portas da família Catalyst 6000, há normalmente quatro limites que são associados às filas de entrada. Geralmente, há dois limiares associados às filas de saída.

Esses limiares também são distribuídos, no contexto do QoS, como um meio de atribuir quadros com diferentes prioridades para tais limiares. À medida que o buffer começa a ser preenchido e os limites são violados, o administrador pode mapear prioridades diferentes para limiares diferentes, indicando para o switch quais quadros devem ser descartados quando um limite é excedido.

Mapeamentos

Nas seções de filas e limiares acima, foi mencionado que o valor de CoS no quadro Ethernet é usado para determinar em que fila colocar o quadro e em que ponto do preenchimento do buffer é um quadro elegível para descarte. Essa é a finalidade dos mapeamentos.

Quando o QoS está configurado na família Catalyst 6000, são habilitados mapeamentos padrão que definem o seguinte:

- em quais quadros de limites com valores de CoS específicos são elegíveis para serem soltos
- em que fila um quadro é colocado (com base em seu valor de CoS)

Enquanto os mapeamentos padrão existirem, eles poderão ser substituídos pelo administrador. Existe um mapeamento para o seguinte:

- Valores de CoS em um quadro de entrada para um valor de DSCP
- Valores de precedência de IP em um quadro de entrada para um valor de DSCP
- Valores de DSCP para um valor de CoS para um quadro de saída
- Valores de CoS para queda de limiares em filas de recebimento
- Valores de CoS para queda de limiares em filas de transmissão

- Valores de marcação DSCP para quadros que excedem as instruções de vigilância
- Valores de CoS para um quadro com um endereço MAC de destino específico

WRED e WRR

WRED e WRR são dois algoritmos extremamente potentes que fazem parte da família Catalyst 6000. Tanto o WRED quanto o WRR usam a marca de prioridade (CoS) dentro de um quadro Ethernet para fornecer gerenciamento de buffer avançado e programação de saída. B

WRED

O WRED é um algoritmo de gerenciamento de buffer empregado pela família Catalyst 6000 para minimizar o impacto da queda do tráfego de alta prioridade em momentos de congestionamento. O WRED é baseado no algoritmo RED.

Para entender o RED e o WRED, revise o conceito de gerenciamento de fluxo do TCP. O gerenciamento de fluxo garante que o remetente do TCP não sobrecarregue a rede. O algoritmo de inicio lento TCP é parte da solução para lidar com isso. Ele determina que quando um fluxo é iniciado, um único pacote é enviado antes de esperar por uma confirmação. Dois pacotes são enviados antes de um ACK ser recebido, aumentando gradualmente o número de pacotes enviados antes do ACK (reconhecimento) ser recebido. Isso continuará até que o fluxo atinja um nível de transmissão (ou seja, envie *x* número de pacotes) que a rede possa lidar sem a carga incorrendo em congestionamento. Se ocorrer congestionamento, o algoritmo de início lento limitará o tamanho da janela (ou seja, o número de pacotes enviados antes de esperar uma confirmação), reduzindo assim o desempenho geral para essa sessão TCP (fluxo).

O RED monitorará uma fila assim que ela começar a ser preenchida. Quando um determinado limite for excedido, os pacotes começarão a ser descartados aleatoriamente. Os fluxos específicos não são tidos em conta; em vez disso, os pacotes aleatórios serão descartados. Esses pacotes podem ser de fluxos de prioridade alta ou baixa. Pacotes descartados podem fazer parte de um único fluxo ou vários fluxos TCP. Se vários fluxos forem afetados, como descrito acima, isso pode ter um impacto considerável em cada tamanho de janela de fluxos.

Diferente do RED, o WRED não é aleatório ao eliminar quadros. O WRED leva em consideração a prioridade das estruturas (no caso da família Catalyst 6000, ele usa o valor CoS). Com WRED, o administrador atribui quadros com certos valores de CoS a limites específicos. Quando esses limiares forem excedidos, os quadros com valores de CoS que estiverem mapeados para esses limiares estarão elegíveis para desconexão. Outros quadros com valores de CoS atribuídos aos thresholds mais altos são mantidos na fila. Esse processo permite que fluxos de prioridade mais alta sejam mantidos intactos, mantendo seus tamanhos maiores de janela intactos e minimizando a latência envolvida na obtenção dos pacotes do remetente para o receptor.

Como você sabe se sua placa de linha suporta WRED? Emita o seguinte comando. Na saída, verifique a seção que indica suporte para WRED nessa porta.

```
Console> show qos info config 2/1
QoS setting in NVRAM:
QoS is enabled
Port 2/1 has 2 transmit queue with 2 drop thresholds (2q2t).
Port 2/1 has 1 receive queue with 4 drop thresholds (1q4t).
```

```
Interface type:vlan-based
ACL attached:
The qos trust type is set to untrusted.
Default CoS = 0
Queue and Threshold Mapping:
Queue Threshold CoS
----- --------- ---------------
1     1         0 1
1     2         2 3
2     1         4 5
2     2         6 7
Rx drop thresholds:
Rx drop thresholds are disabled for untrusted ports.
Queue #  Thresholds - percentage (abs values)
-------  ----------------------------------
1        50% 60% 80% 100%
TX drop thresholds:
Queue #  Thresholds - percentage (abs values)
-------  ----------------------------------
1        40% 100%
2        40% 100%
TX WRED thresholds:
WRED feature is not supported for this port_type.
!-- Look for this. Queue Sizes: Queue # Sizes - percentage (abs values) ------- ----------------
--------------------- 1 80% 2 20% WRR Configuration of ports with speed 1000MBPS: Queue # Ratios
(abs values) ------- ----------------------------------- 1 100 2 255 Console> (enable)
```

Caso o WRED não esteja disponível em uma porta, a porta usará um método de descarte traseiro de gerenciamento de buffer. A queda traseira, como o nome indica, simplesmente descarta os quadros recebidos quando os buffers forem completamente utilizados.

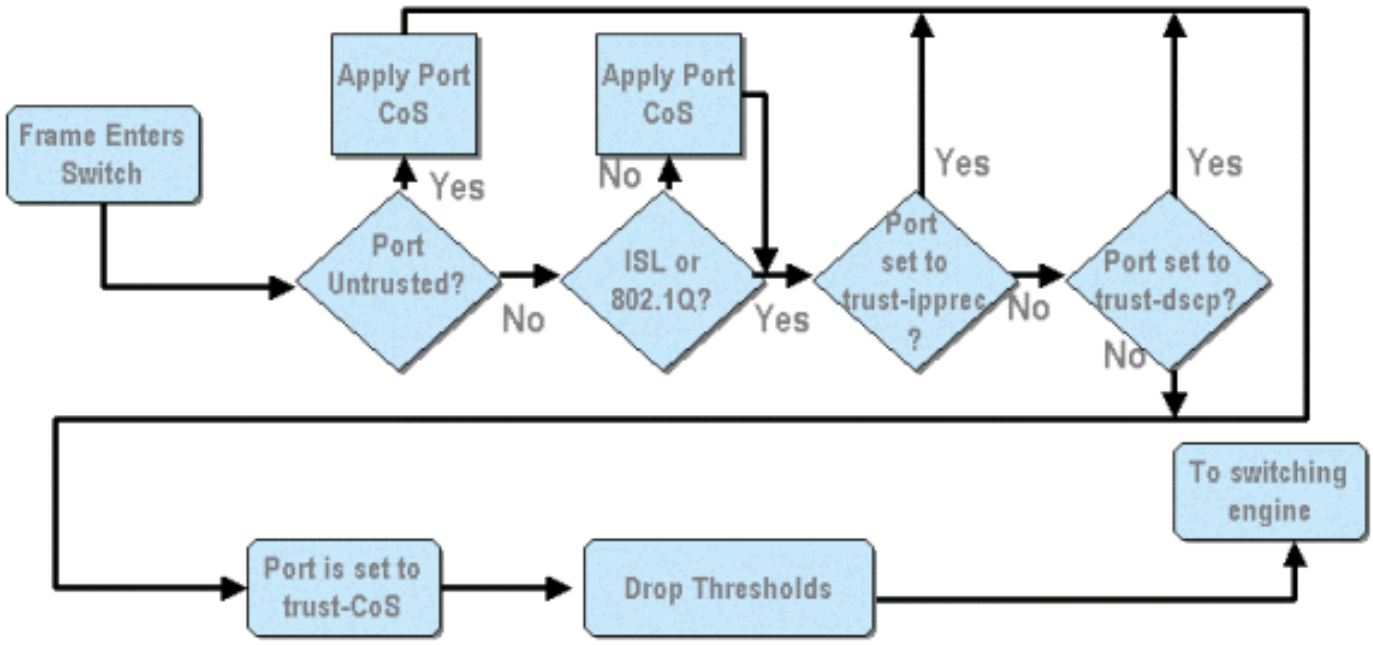
WRR

O WRR é usado para programar o tráfego de saída das filas TX. Um algoritmo de rodízio normal alternará entre as filas TX que enviam um número igual de pacotes de cada fila antes de se mover para a próxima fila. O aspecto ponderado do WRR permite que o algoritmo de programação inspecione um peso que foi atribuído à fila. Isso permite acesso definido a filas para uma parte maior da largura de banda. O algoritmo de agendamento WRR esvaziará mais dados de filas identificadas do que de outras filas, fornecendo assim um viés para filas designadas.

A configuração do WRR e os outros aspectos do que foram descritos acima são explicados nas seções a seguir.

Configurando o QoS com base na porta ASIC no Catalyst 6000 Family

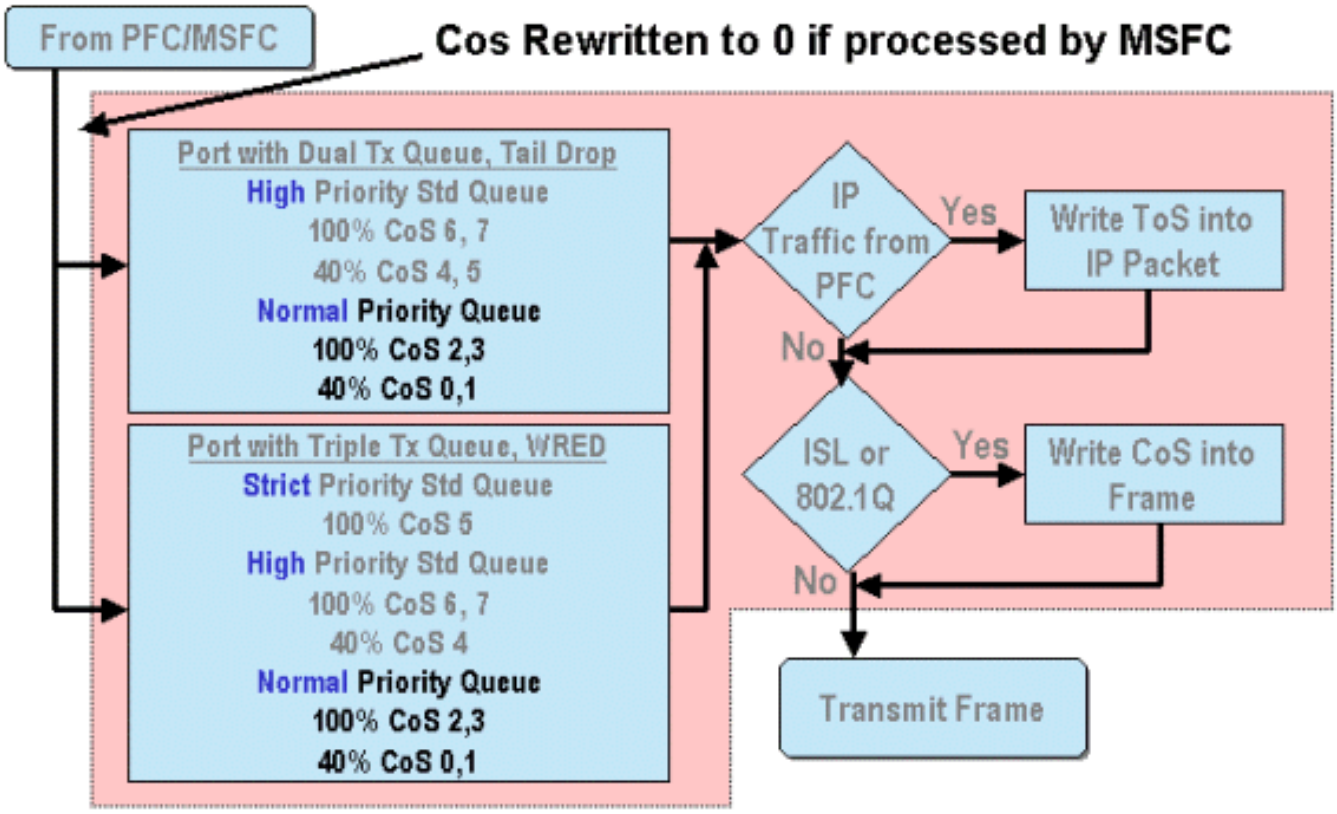
A configuração de QoS instrui o ASIC da porta ou o PFC a executar uma ação de QoS. As seções a seguir examinarão a configuração de QoS para estes dois processos. Na porta ASIC, a configuração QoS afeta os fluxos de tráfego de entrada e saída.



No diagrama acima, pode-se ver que os seguintes processos de configuração de QoS se aplicam:

1. estados confiáveis de portas

2. aplicação de CoS baseado em porta
3. Atribuição de limiar de queda de recebimento
4 Mapas de limiar de queda de CoS para Rx



Quando um quadro é processado por MSFC ou PFC, é passado para a porta de saída ASIC para posterior processamento. Os quadros processados pelo MSFC terão seus valores de CoS redefinidos como zero. Isso deve ser levado em consideração para o processamento de QoS nas portas externas.

O diagrama acima mostra o processamento de QoS executado pelo ASIC de porta para tráfego de saída. Alguns dos processos acionados no processamento de saída QoS incluem o seguinte:

1. Atribuições de queda traseira de TX e limiar de WRED
2. CoS para queda traseira de TX e mapas WRED

Além disso, não mostrado no diagrama acima, é o processo de reatribuir o CoS ao quadro de saída usando um mapa de DSCP para CoS.

As seções a seguir examinam os recursos de configuração de QoS dos ASICs baseados em portas com mais detalhes.

Note: Um ponto importante a ser observado é que quando os comandos QoS são chamados usando CatOS, eles geralmente se aplicam a todas as portas com o tipo de fila especificado. Por exemplo, se um limite de queda WRED for aplicado a portas com tipo de fila 1p2q2t, esse limite de queda WRED será aplicado a todas as portas em todas as placas de linha que suportam esse tipo de fila. Com o Cat IOS, os comandos do QoS são geralmente aplicados no nível da interface.

Habilitando o QoS

Antes que qualquer configuração de QOS possa ocorrer na família Catalyst 6000, a QoS deve primeiro ser ativada no switch. Para fazer isso, emita o seguinte comando:

CatOS

```
Console> (enable) set qos enable
!-- QoS is enabled. Console> (enable)
```

Cisco IOS integrado (modo nativo)

```
Cat6500(config)# mls qos
```

Quando a QoS é habilitada na família Catalyst 6000, o switch definirá uma série de padrões de QoS para o switch. Esses padrões incluem as seguintes configurações:

QoS Feature	Default setting
Trust state of each port	Un-trusted
Receive Queue drop threshold percentages	Threshold 1 – 50% Threshold 2 – 60% Threshold 3 – 80% Threshold 4 – 100%
Transmit Queue drop threshold percentages	Low priority queue threshold 1 – 80% Low priority queue threshold 2 – 100% High priority queue threshold 1 – 80% High priority queue threshold 2 – 100%
CoS value to Drop threshold mapping	Receive queue 1/drop threshold 1: CoS 0 and 1 Transmit queue 1/drop threshold 1: CoS 0 and 1 Receive queue 1/drop threshold 2: CoS 2 and 3 Transmit queue 1/drop threshold 2: CoS 2 and 3 Receive queue 1/drop threshold 3: CoS 4 and 5 Transmit queue 2/drop threshold 1: CoS 4 and 5 Receive queue 1/drop threshold 4: CoS 6 and 7
	Transmit queue 2/drop threshold 2: CoS 6 and 7
CoS to DSCP Mapping (DSCP set from CoS value)	CoS 0 = DSCP 0 CoS 1 = DSCP 8 CoS 2 = DSCP 16 CoS 3 = DSCP 24 CoS 4 = DSCP 32 CoS 5 = DSCP 40 CoS 6 = DSCP 48 CoS 7 = DSCP 56
IP Precedence to DSCP Map (DSCP set from IP Precedence value)	IP precedence 0 = DSCP 0 IP precedence 1 = DSCP 8 IP precedence 2 = DSCP 16 IP precedence 3 = DSCP 24 IP precedence 4 = DSCP 32 IP precedence 5 = DSCP 40 IP precedence 6 = DSCP 48 IP precedence 7 = DSCP 56
DSCP to CoS map (CoS set from DSCP values)	DSCP 0-7 = CoS 0 DSCP 8-15 = CoS 1 DSCP 16-23 = CoS 2 DSCP 24-31 = CoS 3 DSCP 32-39 = CoS 4 DSCP 40-47 = CoS 5 DSCP 48-55 = CoS 6 DSCP 56-63 = CoS 7

Portas confiáveis e não confiáveis

Qualquer porta dada na família Catalyst 6000 pode ser configurada como confiável ou não confiável. O estado de confiança da porta determina como marca, classifica e agenda o quadro à medida que ele transita pelo switch. Por padrão, todas as portas estão no estado não confiável.

Portas Não-Confiáveis (Configuração Padrão de Portas)

Caso a porta seja configurada como não confiável, um quadro, depois de entrar inicialmente na porta, terá seus valores CoS e ToS zerados pela porta ASIC. Isso significa que o quando receberá o serviço de menor prioridade no caminho pelo Switch.

Como alternativa, o administrador pode redefinir o valor de CoS de qualquer quadro Ethernet que entre uma porta não confiável em um valor predeterminado. A configuração disso será discutida em uma seção posterior.

Definir a porta como não-confiável instruirá o Switch a não realizar nenhuma fuga de congestionamento. Evitar congestionamento é o método usado para descartar quadros com base em seus valores de CoS quando eles excedem os limiares definidos para essa fila. Todos os quadros que entram nessa porta serão igualmente qualificados para serem descartados quando os buffers atingirem 100%.

No CatOS, uma porta 10/100 ou GE pode ser configurada como não confiável emitindo o seguinte comando:

CatOS

```
Console> (enable) set port qos 3/16 trust untrusted
!-- Port 3/16 qos set to untrusted. Console> (enable)
```

Esse comando configura a porta 16 do módulo 3 como não confiável.

Note: Para o Cisco IOS integrado (modo nativo), o software atualmente suporta apenas a definição de confiança para portas GE.

Cisco IOS integrado (modo nativo)

```
Cat6500(config)# interface gigabitethernet 1/1
Cat6500(config-if)# no mls qos trust
```

No exemplo acima, inserimos a configuração da interface e aplicamos a forma **no** do comando para definir a porta como não confiável, já que é o IOS.

Portas Confiáveis

Às vezes, os quadros Ethernet que entram em um switch terão uma configuração de CoS ou ToS que o administrador deseja que o switch mantenha enquanto o quadro transita pelo switch. Para esse tráfego, o administrador pode definir o estado de confiança de uma porta onde esse tráfego entra no switch como confiável.

Como mencionado anteriormente, o switch usa internamente um valor de DSCP para atribuir um nível de serviço predeterminado a esse quadro. À medida que um quadro entra em uma porta confiável, o administrador pode configurar a porta para examinar o valor de CoS, precedência de IP ou DSCP existente para definir o valor de DSCP interno. Como alternativa, o administrador pode definir um DSCP predefinido para cada pacote que entra na porta.

A configuração do estado de confiança de uma porta como confiável pode ser alcançada emitindo o seguinte comando:

CatOS

```
Console> (enable) set port qos 3/16 trust trust-cos
!-- Port 3/16 qos set to trust-COs Console> (enable)
```

Esse comando é aplicável na placa WS-X6548-RJ45 e define o estado de confiança da porta 3/16 como confiável. O Switch usará o valor CoS definido no quadro de entrada para definir o DSCP interno. O DSCP é derivado de um mapa padrão criado quando a QoS foi habilitada no switch ou, alternativamente, de um mapa definido pelo administrador. Em vez da palavra-chave trust-COs, o administrador também pode usar as palavras-chave trust-dscp ou trust-ipprec.

Em placas de ingresso 10/100 anteriores (WS-X6348-RJ45 e WS-X6248-RJ45), a confiança de portas precisa ser definida emitindo o comando set qos acl. Neste comando, um estado confiável pode ser atribuído por um subparâmetro do comando **set qos acl**. A configuração de trust CoS não é suportada para portas dessas placas de linha, conforme descrito abaixo.

```
Console> (enable) set port qos 4/1 trust trust-COs
Trust type trust-COs not supported on this port.
!-- Trust-COs not supported, use acl instead. Rx thresholds are enabled on port 4/1. !-- Need to
turn on input queue scheduling. Port 4/1 qos set to untrusted. !-- Trust-COs not supported, so
port is set to untrusted.
```

O comando acima indica que é necessário ativar o agendamento da fila de entrada. Portanto, para portas 10/100 em placas de linha WS-X6248-RJ45 e WS-X6348-RJ45, o comando set port qos x/y trust trust-COs deve estar ainda configurado, apesar de que o ALC deve ser utilizado para configurar estados de confiança.

Com o Cisco IOS integrado (modo nativo), a configuração de confiança pode ser executada em uma interface GE e em portas 10/100 na nova placa de linha WS-X6548-RJ45.

Cisco IOS integrado (modo nativo)

```
Cat6500(config)# interface gigabitethernet 5/4
Cat6500(config-if)# mls qos trust ip-precedence
Cat6500(config-if)#
```

Este exemplo configura o estado de confiança da porta GE 5/4 como confiável. O valor de precedência de IP do quadro será usado para derivar o valor do DSCP.

Classificação de entrada e configuração de CoS baseado em porta

Na entrada em uma porta de switch, um quadro Ethernet pode ter seu CoS alterado se atender a um dos dois critérios a seguir:

1. a porta está configurada como não confiável, ou
2. a estrutura de Ethernet não tem um valor COS existente já configurado.

Se desejar reconfigurar o CoS de um quadro Ethernet de entrada, você deve emitir o seguinte comando:

CatOS

```
Console> (enable) set port qos 3/16 cos 3
!-- Port 3/16 qos set to 3. Console> (enable)
```

Esse comando configura os COs de quadros Ethernet de entrada na porta 16 do módulo 3 para um valor de 3 quando um quadro não marcado chega ou quando a porta está configurada como não confiável.

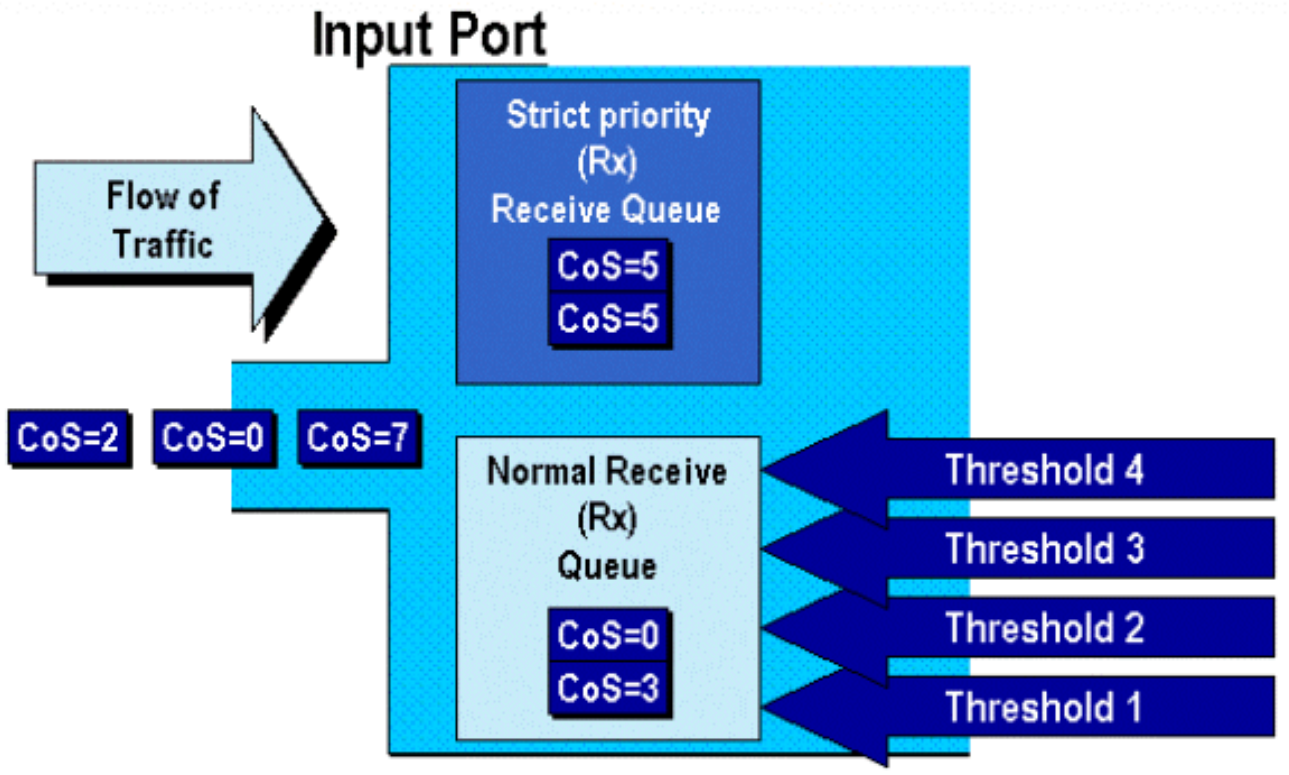
Cisco IOS integrado (modo nativo)

```
Cat6500(config)# interface fastethernet 5/13
Cat6500(config-if)# mls qos COs 4
Cat6500(config-if)#
```

Esse comando define os COs dos quadros Ethernet de entrada na porta 13 no módulo 5 para um valor de 4 quando um quadro não marcado chega ou se a porta está definida como não confiável.

Configure Rx Drop Thresholds

On ingress to the Switch port, the frame will be placed into a Rx queue. Para evitar o excesso de buffers, a porta ASIC implementa quatro limiares em cada fila Rx e usa esses limiares para identificar quadros que podem ser descartados uma vez que esses limiares são excedidos. O ASIC de porta irá utilizar o valor de COs de conjunto de quadros para identificar quais quadros podem ser derrubados quando um limiar é excedido. Esse recurso permite que os quadros com prioridade mais elevada permaneçam no buffer por mais tempo quando ocorre congestionamento.



Como mostrado no diagrama acima, os quadros chegam e são colocados na fila. À medida que a fila começa a ser preenchida, os limiares são monitorados pelo ASIC da porta. Quando um limiar é rompido, estruturas com valores de CO identificados pelo administrador são descartadas

aleatoriamente da fila. Os mapeamentos de limiar padrão para uma fila 1a4t (encontrados nas placas de ingresso WS-X6248-RJ45 e WS-X6348-RJ45) são os seguintes:

- o limiar 1 é definido para 50% e os valores COs 0 e 1 são mapeados para este limiar
- o limiar 2 é definido para 60% e os valores COs 2 e 3 são mapeados para este limiar
- o limiar 3 é definido para 80% e os valores COs 4 e 5 são mapeados para este limiar
- o limiar 4 é definido para 100% e os valores COs 6 e 7 são mapeados para este limiar

Para uma fila 1P1q4t (encontrada em portas GE), os mapeamentos padrão são os seguintes:

- o limiar 1 é definido para 50% e os valores COs 0 e 1 são mapeados para este limiar
- o limiar 2 é definido para 60% e os valores COs 2 e 3 são mapeados para este limiar
- o limiar 3 é definido como 80% e os valores de CO 4 são mapeados para este limiar
- o limiar 4 é definido para 100% e os valores COs 6 e 7 são mapeados para este limiar
- O valor de COs 5 é mapeado para a fila de prioridade estrita

Para uma 1p1q0t (encontrada em portas 10/100 na placa de linha WS-X6548-RJ45), os mapeamentos padrão são os seguintes:

- Os quadros com COs 5 vão para a fila SP Rx (fila 2), onde o switch descarta os quadros de entrada somente quando o buffer da fila de recepção do SP estiver 100% cheio.
- Os quadros com COs 0, 1, 2, 3, 4, 6 ou 7 vão para a fila Rx padrão. O switch descarta quadros de entrada quando o buffer de fila Rx está 100% cheio.

Esses limiares de queda podem ser alterados pelo administrador. Além disso, os valores de COs padrão que são mapeados para cada limiar também podem ser alterados. Diferentes placas de linha implementam diferentes implementações de fila Rx. Um resumo dos tipos de fila é mostrado abaixo.

CatOS

```
Console> (enable) set qos drop-threshold 1q4t rx queue 1 20 40 75 100
!-- Rx drop thresholds for queue 1 set at 20%, 40%, 75%, and 100%. Console> (enable)
```

Esse comando define os limites de queda de recebimento de todas as portas de entrada com uma fila e quatro limiares (significa 1q4t) para 20%, 40%, 75% e 100%.

O comando emitido no Integrated Cisco IOS (Modo Nativo) é mostrado a seguir.

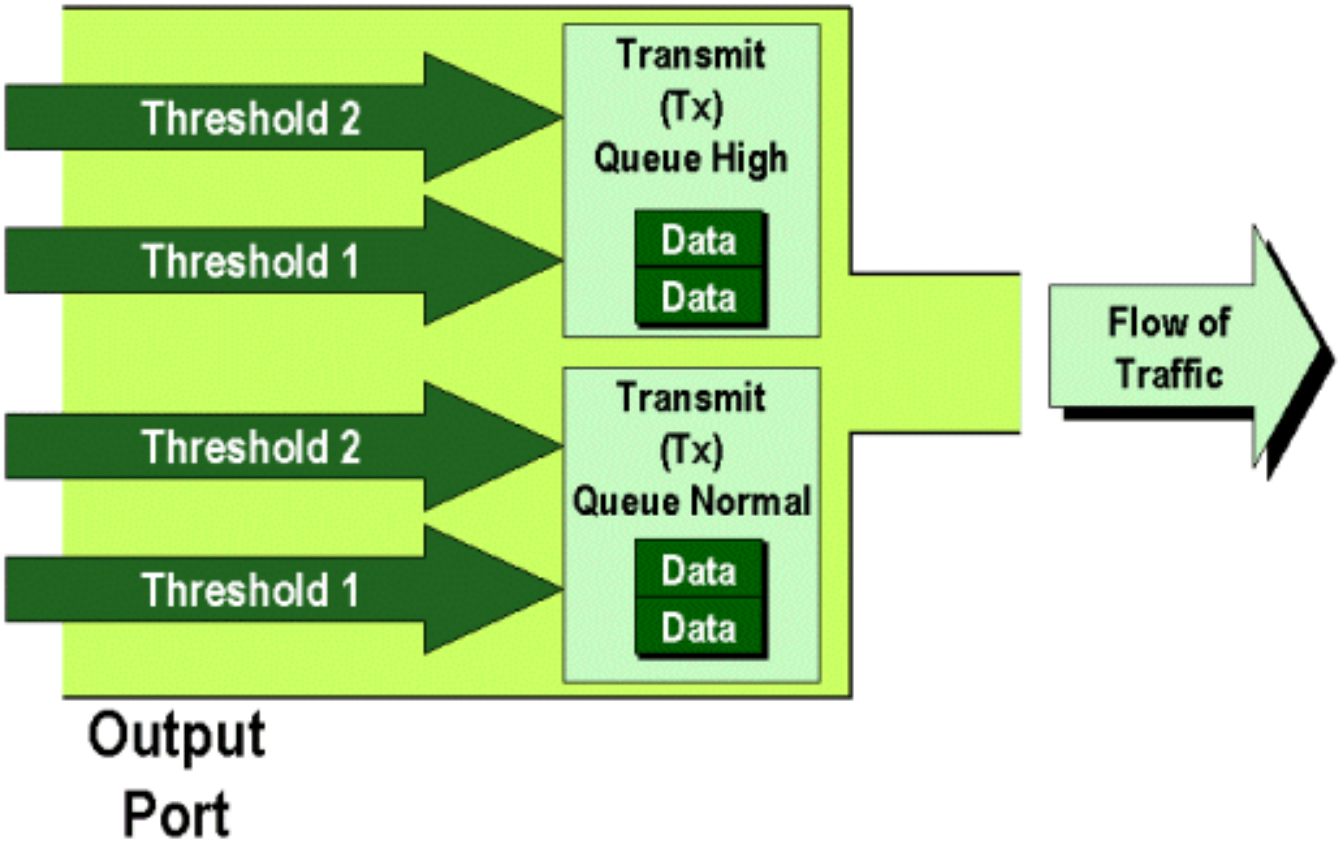
Cisco IOS integrado (modo nativo)

```
Cat6500(config-if)# wrr-queue threshold 1 40 50
Cat6500(config-if)# wrr-queue threshold 2 60 100

!-- Configures the 4 thresholds for a 1q4t rx queue and. Cat6500(config-if)# rcv-queue threshold
1 60 75 85 100

!-- Configures for a 1p1q4t rx queue, which applies to !-- the new WS-X6548-RJ45 10/100 line
card.
```

Os limiares de queda de Rx devem ser habilitados pelo administrador. Atualmente, o comando **set port qos x/y trust trust-COs** deve ser usado para ativar os limiares de queda de Rx (onde *x* é o número do módulo e *y* é a porta nesse módulo).

Configuração de limiares TX Drop

Em uma porta de saída, a porta terá dois limiares de TX que são usados como parte do mecanismo de prevenção de congestionamento, a fila 1 e a fila 2. A fila 1 é indicada como a fila padrão de baixa prioridade, e a fila 2 é indicada como a fila padrão de alta prioridade. Dependendo das placas de linha usadas, elas empregarão um tail drop ou um algoritmo de gerenciamento de limiar WRED. Ambos os algoritmos empregam dois limiares para cada fila de TX.



O administrador pode configurar manualmente os limiares da seguinte maneira:

CatOS

```
Console> (enable) set qos drop-threshold 2q2t TX queue 1 40 100
!-- TX drop thresholds for queue 1 set at 40% and 100%. Console> (enable)
```

Esse comando define os limiares de queda de TX para a fila 1 para todas as portas de saída com duas filas e dois limiares (denota 2q2t) como 40% e 100%.

```
Console> (enable) set qos wred 1p2q2t TX queue 1  60 100
!-- WRED thresholds for queue 1 set at 60% 100% on all WRED-capable 1p2q2t ports. Console>
(enable)
```

Esse comando configura os limiares de queda do WRED para fila 1 para todas as portas de saída com uma fila SP, duas filas normais e dois limiares (indica 1p2q2t) para 60% e 100%. A fila 1 é definida como a fila de prioridade normal baixa e apresenta a prioridade mais baixa. A fila 2 é a fila normal de alta prioridade e tem uma prioridade mais alta que a fila 1. A fila 3 é a fila do SP e é atendida antes de todas as outras filas nessa porta.

O comando equivalente emitido no Cisco IOS integrado (modo nativo) é mostrado abaixo.

Cisco IOS integrado (modo nativo)

```
Cat6500(config-if)# wrr-queue random-detect max-threshold  1 40 100
Cat6500(config-if)#
```

Isto define os limiares de queda WRED de uma porta 1p2q2t para a fila 1 em 40% do limiar 1 (TX) e 100% do limiar 2 (TX).

O WRED também poderá ser desabilitado se necessário no Cisco IOS Integrado (Modo Nativo). O método usado para fazer isso é usar a forma **n"** do comando. Um exemplo de como desativar o WRED é mostrado a seguir:

Cisco IOS integrado (modo nativo)

```
Cat6500(config-if)# no wrr-queue random-detect queue_id
```

Mapeando o endereço MAC para valores de COs

Além de definir COs com base em uma definição de porta global, o switch permite que o administrador defina valores de COs com base no endereço MAC de destino e na ID da VLAN. Isso permite que os quadros destinados a destinos específicos sejam marcados com um valor de CO predeterminado. Esta configuração pode ser feita emitindo o comando a seguir:

CatOS

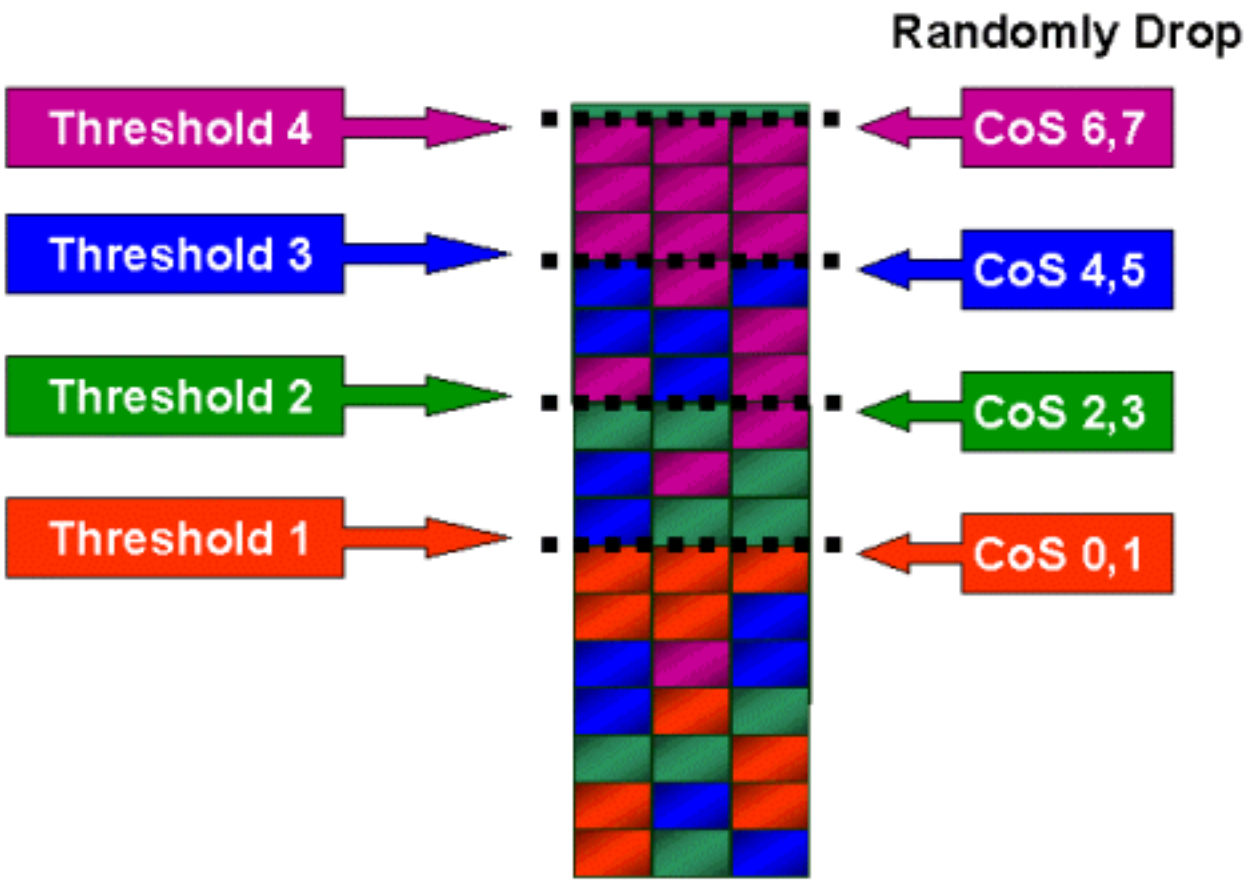
```
Console> (enable) set qos Mac-COs 00-00-0c-33-2a-4e 200 5
!-- COs 5 is assigned to 00-00-0c-33-2a-4e VLAN 200. Console> (enable)
```

Este comando configura um COs de 5 para qualquer quadro cujo endereço MAC de destino seja 00-00-0c-33-2a-4e que tenha vindo da VLAN 200.

Não há nenhum comando equivalente no Cisco IOS integrado (modo nativo). Isso ocorre porque este comando é suportado apenas quando não há uma PFC presente e o Integrated Cisco IOS (modo Nativo) requer uma PFC para funcionar.

Mapeamento de COs para limites

Depois que os limites tiverem sido configurados, o administrador poderá atribuir valores de COs a esses limiares, de modo que quando o limite tiver sido excedido, os quadros com valores específicos de COs poderão ser descartados. Normalmente, o administrador atribuirá quadros de prioridade inferior aos limiares inferiores, mantendo assim o tráfego de prioridade superior na fila, caso ocorra congestionamento.



A figura acima mostra uma fila de entrada com quatro limiares e como os valores de COs foram atribuídos a cada limite.

A seguinte saída apresenta como os valores de COs podem ser mapeados para limiares:

CatOS

```
Console> (enable) set qos map 2q2t 1 1 COs 0 1
!-- QoS TX priority queue and threshold mapped to COs successfully. Console> (enable)
```

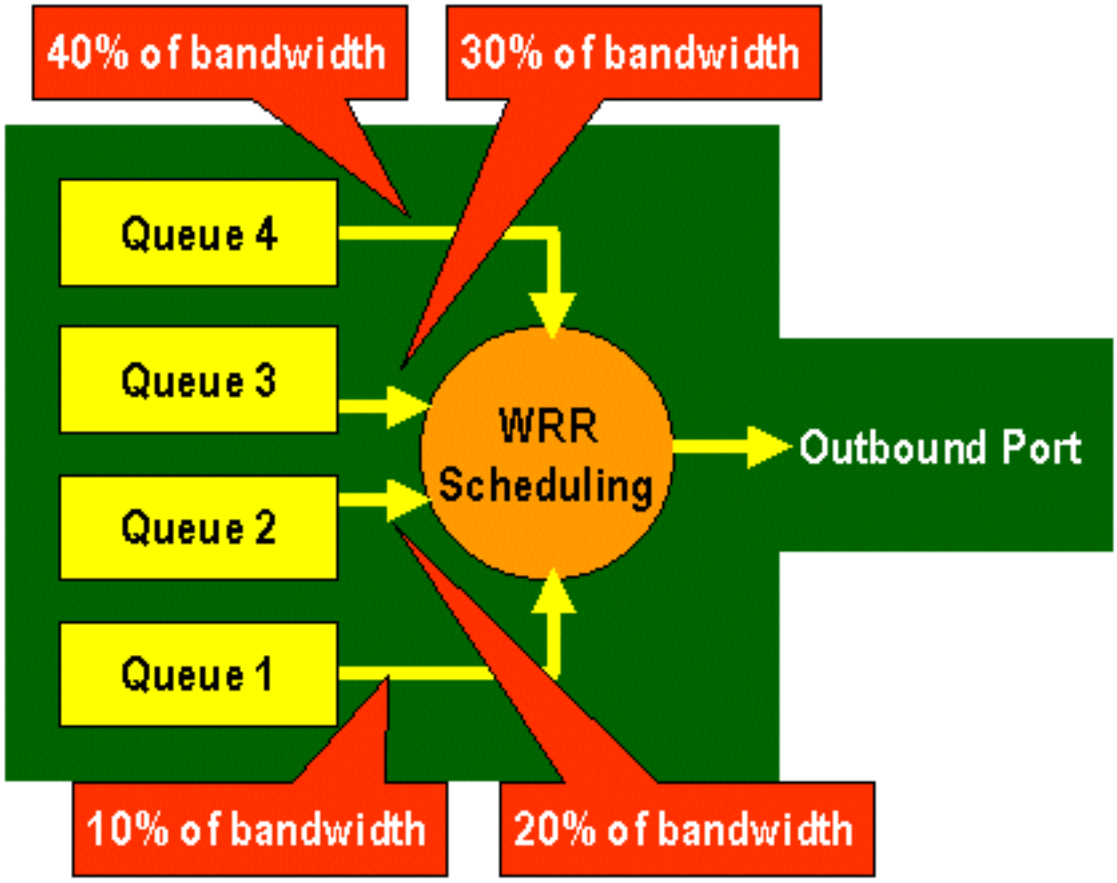
Esse comando atribui valores COs de 0 e 1 à fila 1, limiar 1. O comando equivalente no Cisco IOS integrado (modo nativo) é mostrado abaixo.

Cisco IOS integrado (modo nativo)

```
Cat6500(config-if)# wrr-queue COs-map  1  1  0  1
Cat6500(config-if)#
```

Configuração da largura de banda em filas TX

Se uma estrutura for colocada em uma fila de saída, ela será transmitida com o uso de um algoritmo output-scheduling. O processo do programador de saída usa o WRR para transmitir quadros a partir das filas de saída. Dependendo do hardware da placa de linha sendo usado, há duas, três ou quatro filas de transmissão por porta.



Nas placas de ingresso WS-X6248 e WS-X6348 (com estruturas de fila 2q2t), duas filas TX são usadas pelo mecanismo WRR para programação. Nas placas de linha WS-X6548 (com uma estrutura de fila 1p3q1t) há quatro filas TX. Dessas quatro filas TX, três filas TX são atendidas pelo algoritmo WRR (a última fila TX é uma fila SP). Em placas de linha GE, há três filas TX (usando uma estrutura de fila 1p2q2t); uma dessas filas é uma fila SP, portanto, o algoritmo WRR atende apenas duas filas TX.

Geralmente, o administrador atribuirá um peso à fila TX. O WRR funciona verificando o peso atribuído à fila de portas, que é usado internamente pelo Switch para determinar a quantidade de tráfego a ser transmitida, antes de passar para a próxima fila. Um valor de ponderação entre 1 e 255 pode ser atribuído a cada fila de portas.

CatOS

```
Console> (enable) set qos wrr 2q2t 40 80
!-- QoS wrr ratio set successfully. Console> (enable)
```

Esse comando atribui uma ponderação de 40 à fila 1 e 80 à fila 2. Isso significa efetivamente uma proporção de dois para um (80 a 40 = 2 a 1) da largura de banda atribuída entre as duas filas. Este comando tem efeito em todas as portas com duas filas e dois limiares.

O comando equivalente emitido no Cisco IOS integrado (modo nativo) é mostrado abaixo.

Cisco IOS integrado (modo nativo)

```
Cat6500(config-if)# wrr-queue bandwidth 1 3
```

```
Cat6500(config-if)#
```

Os dados acima representam uma proporção de três para um entre as duas filas. Você observará que a versão do Cat IOS desse comando se aplica somente a uma interface específica.

DSCP para mapeamento de COs

Depois de colocado o frame na porta de saída, a porta ASCII usará os COs atribuídos para executar fuga de congestionamento (ou seja, WRED) e também utilizará os COs para determinar a programação do frame (ou seja, a transmissão do frame). Nesse ponto, o switch usará um mapa padrão para retornar o DSCP atribuído a um valor de COs. Este mapa predefinido é apresentado nesta tabela.

Como alternativa, o administrador pode criar um mapa que será usado pelo switch para pegar o valor de DSCP interno atribuído e criar um novo valor de COs para o quadro. Exemplos de como você usaria o CatOS e o Integrated Cisco IOS (modo nativo) para fazer isso são mostrados abaixo.

CatOS

```
Console> (enable) set qos dscp-cos--map 20-30:5 10-15:3 45-52:7
!-- QoS dscp-cos-map set successfully. Console> (enable)
```

O comando acima mapeia valores de DSCP de 20 a 30 para um valor de COs de 5, valores de DSCP de 10 a 15 para um COs de 3, e valores de DSCP de 45 a 52 para um valor de COs de 7. Todos os outros valores de DSCP usam o mapa padrão criado quando a QoS foi habilitada no switch.

O comando equivalente emitido no Cisco IOS integrado (modo nativo) é mostrado abaixo.

Cisco IOS integrado (modo nativo)

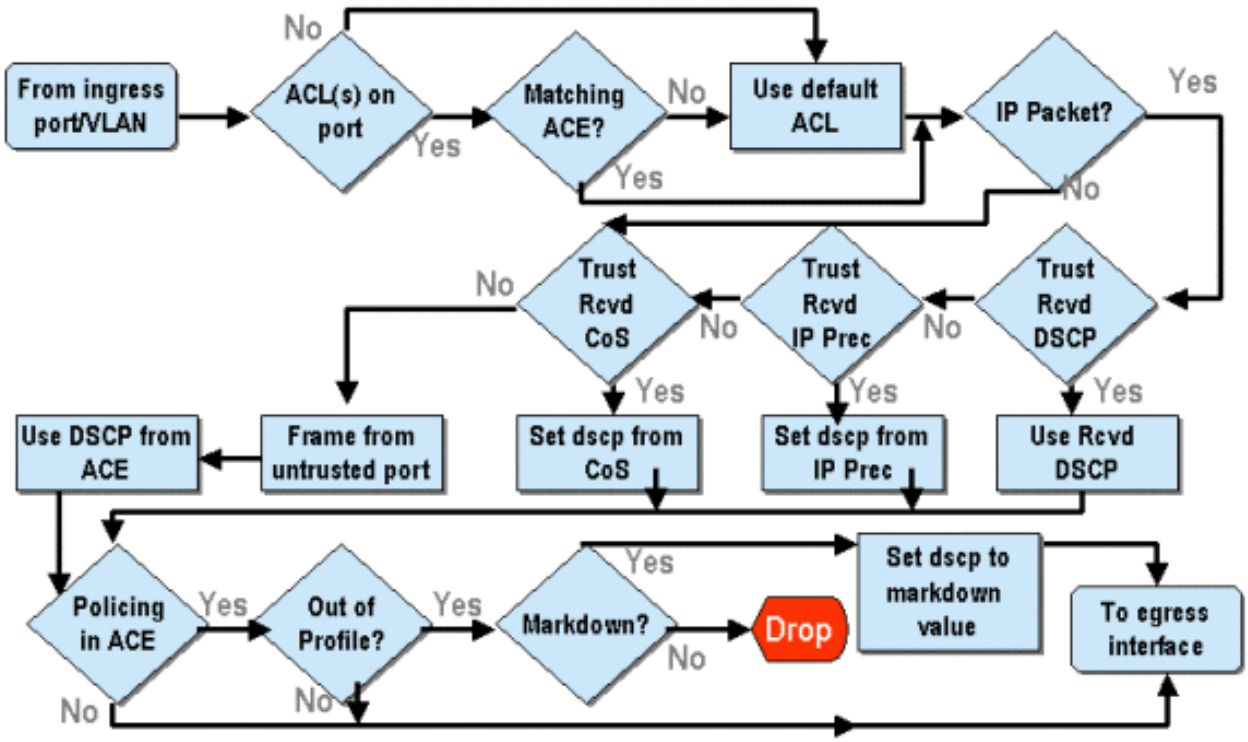
```
Cat6500(config)# mls qos map dscp-cos 20 30 40 50 52 10 1 to 3
Cat6500(config)#
```

Configura os valores DSCP de 20, 30, 40, 50, 52, 10 e 1 para um valor de 3 de COs.

Classificação e vigilância com o PFC

O PFC suporta a classificação e a vigilância de quadros. A classificação pode usar uma ACL para atribuir (marcar) um quadro de entrada com uma prioridade (DSCP). O policiamento permite que um fluxo de tráfego seja limitado a uma certa quantidade de largura de banda.

As seções a seguir descreverão esses recursos no PFC da perspectiva do CatOS e das plataformas do SO Integrated Cisco IOS (modo nativo). Os processos aplicados pelo PFC são mostrados no diagrama a seguir:



Configurar políticas na família Catalyst 6000 com CatOS

A função de vigilância é dividida em duas seções, uma para CatOS e uma para Cisco IOS integrado (modo nativo). Ambos obtêm o mesmo resultado final, mas são configurados e implementados de maneiras diferentes.

Vigilância

A PFC suporta a capacidade de limitar (ou policiar) o tráfego de entrada para o switch e pode reduzir o fluxo de tráfego para um limite predefinido. O tráfego excedente a esse limite pode ser descartado ou ter o valor DSCP marcado no quadro como menor.

A limitação da taxa de saída (saída) não é suportada atualmente no PFC1 ou no PFC2. Isso será adicionado em uma nova revisão do PFC planejada para o segundo semestre de 2002, que suportará a vigilância de saída (ou saída).

O policiamento é suportado no CatOS e no novo Cisco IOS Integrado (Modo Nativo), embora a configuração desses recursos seja muito diferente. As seguintes seções descreverão a configuração de vigilância nas duas plataformas de OS.

Agregados e Microfluxos (CatOS)

Agregados e Microfluxos são termos usados para definir o escopo de policiamento que a PFC executa.

Um microfluxo define a vigilância de um único fluxo. Um fluxo é definido por uma sessão com um endereço MAC SA/DA exclusivo, endereço IP SA/DA e números de porta TCP/UDP. Para cada novo fluxo iniciado por uma porta de uma VLAN, o microfluxo pode ser usado para limitar a quantidade de dados recebidos para esse fluxo pelo switch. Na definição de microfluxo, os pacotes que excedem o limite de taxa prescrito podem ser descartados ou ter seu valor de DSCP marcado como inativo.

Semelhante a um microfluxo, um agregado pode ser usado para limitar a taxa de tráfego. No

entanto, a taxa agregada se aplica a todo o tráfego de entrada em uma porta ou VLAN que corresponda a uma ACL de QoS especificada. Você pode exibir o agregado como a vigilância do tráfego cumulativo que corresponde ao perfil na entrada de controle de acesso (ACE).

Tanto o agregado quanto o microfluxo definem a quantidade de tráfego que pode ser aceita no switch. Um agregado e um microfluxo podem ser atribuídos ao mesmo tempo a uma porta ou a uma VLAN.

Ao definir microfluxos, pode-se definir até 63 deles e até 1023 agregados.

Entradas de controle de acesso e ACLs de QoS (CatOS)

Uma ACL de QoS consiste em uma lista de ACEs que definem um conjunto de regras de QoS que a PFC usa para processar quadros de entrada. Os Aces são semelhantes a uma RACL (Router Access Control List, lista de controle de acesso do roteador). O ACE define critérios de classificação, marcação e vigilância para um quadro de entrada. Se um quadro de entrada corresponder aos critérios definidos na ACE, o mecanismo de QoS processará o quadro (conforme considerado pela ACE).

Todo o processamento de QoS é feito no hardware, portanto, ativar a vigilância de QoS não afeta o desempenho do switch.

O PFC2 suporta atualmente até 500 ACLs e essas ACLs podem consistir em até 32000 Aces (no total). Os números ACE reais dependerão de outros serviços definidos e da memória disponível na PFC.

Existem três tipos de ACEs que podem ser definidos. São eles: IP, IPX e MAC. O acesso IP e IPX inspecionam as informações do cabeçalho L3, enquanto o acesso baseado em MAC inspeciona apenas as informações do cabeçalho L2. Também deve ser observado que o MAC Aces só pode ser aplicado ao tráfego não IP e não IPX.

Criando regras de vigilância

O processo de criação de uma regra de policiamento envolve a criação de um agregado (ou microfluxo) e, em seguida, o mapeamento desse agregado (ou microfluxo) para uma ACE.

Se, por exemplo, o requisito era limitar todo o tráfego IP de entrada na porta 5/3 a um máximo de 20 MB, as duas etapas mencionadas acima devem ser configuradas.

Primeiro, o exemplo solicita que todo o tráfego IP de entrada seja limitado. Isso implica que um vigilante agregado deve ser definido. Um exemplo disso pode ser o seguinte:

```
Console> (enable) set qos policer aggregate test-flow rate 20000 burst 13 policed-dscp
!-- Hardware programming in progress !-- QoS policer for aggregate test-flow created
successfully. Console> (enable)
```

Criamos um agregado chamado de fluxo de teste. Ele define uma taxa de 20.000 KBPS (20 MBPS) e uma intermitência de 13. A palavra-chave policed-dscp indica que qualquer dado que exceda esta política terá seu valor de DSCP marcado como inativo conforme especificado em um mapa de marcação DSCP (existe um padrão ou isso pode ser modificado pelo administrador). Uma alternativa para usar a palavra-chave policed-dscp é usar a palavra-chave drop. A palavra-chave drop simplesmente descartará todo o tráfego fora de perfil (tráfego que fica fora do valor de intermitência distribuído).

A facilidade de vigilância funciona em um esquema de vazamento de token bucket, no sentido de que você define uma intermitência (quantidade de dados em bits por segundo que você aceitará em um dado intervalo (fixo) de tempo) e, depois, a taxa (definida como a quantidade de dados que você retirará daquele bucket em um único segundo). Todos os dados que sobrecarregam esse bucket são descartados ou têm seu DSCP marcado como inativo. O período de tempo (ou intervalo) especificado mencionado acima é de 0,00025 segundos (ou 1/4000 segundo) e é fixo (isto é, não é possível utilizar nenhum comando de configuração para alterar esse número).

O número 13 do exemplo acima representa um bucket que aceitará até 13.000 bits de dados a cada 1/4000 de um segundo. Isso se refere a 52 MB por segundo (13K * (1 / 0.00025) ou 13K * 4000). Você deve sempre verificar se a intermitência está configurada para ser igual ou superior à taxa na qual deseja enviar dados. Em outras palavras, a intermitência deve ser maior ou igual à quantidade mínima de dados que você deseja transmitir para um determinado período. Se a intermitência resultar em um número menor para o que você especificou como sua taxa, o limite de taxa será igual à intermitência. Em outras palavras, se você definir uma taxa de 20 MBPS e uma intermitência que calcule em 15 MBPS, sua taxa só chegará a 15 MBPS. A próxima pergunta que você pode ter é por que 13?. Lembre-se de que o burst define a profundidade do token bucket, ou, em outras palavras, a profundidade do bucket utilizado para receber os dados que chegam a cada 1/4000 de segundo. Portanto, a intermitência pode ser qualquer número suportado em uma taxa de dados de chegada maior ou igual a 20 MB por segundo. O burst mínimo que poderia ser usado para um limite de taxa de 20 MB é 20000/4000 = 5.

Durante o processamento do vigilante, o algoritmo de vigilância começa preenchendo o token bucket com um complemento completo de tokens. O número de tokens é igual ao valor do burst. Então, se o valor de intermitência é 13, o número de tokens no balde é igual a 13.000. Para cada 1/4000 de um segundo, o algoritmo de vigilância enviará uma quantidade de dados igual à taxa definida dividida por 4000. Para cada bit (dígito binário) dos dados enviados, ele consome um token do bucket. No final do intervalo, ele reabastecerá o bucket com um novo conjunto de tokens. O número de tokens que ele substitui é definido pela taxa / 4000. Considere o exemplo acima para entender isso:

```
Console> (enable) set qos policer aggregate test-flow rate 20000 burst 13
```

Suponha que esta seja uma porta de 100 MBPS e que estejamos enviando um fluxo constante de 100 MBPS para a porta. Sabemos que isso equivalerá a uma taxa de entrada de 100.000.000 bits por segundo. Os parâmetros aqui são uma taxa de 20000 e burst de 13. No intervalo de tempo t0, há um complemento completo de tokens no bucket (que é 13.000). No intervalo de tempo t0, teremos o primeiro conjunto de dados chegando à porta. Para esse intervalo de tempo, a taxa de chegada será de 100.000.000 / 4.000 = 25.000 bits por segundo. Como nosso token bucket tem apenas uma profundidade de 13.000 tokens, apenas 13.000 bits dos 25.000 bits que chegam à porta nesse intervalo são elegíveis para envio e 12.000 bits são descartados.

A taxa especificada define uma taxa de encaminhamento de 20.000.000 bits por segundo, o que equivale a 5.000 bits enviados por 1/4.000 intervalo. Para cada 5.000 bits enviados, há 5.000 tokens consumidos. No intervalo de tempo T1, chegam outros 25.000 bits de dados, mas o bucket ca 12.000 bits. O bucket é reposto com tokens definidos como a taxa / 4000 (que equivale a 5.000 novos tokens). O algoritmo emite, em seguida, o próximo complemento de dados, que se iguala a outros 5.000 bits de dados (isso consome outros 5.000 tokens), e assim por diante, em cada intervalo.

Essencialmente, todos os dados que chegam além da profundidade do bucket (burst definido) são descartados. Os dados restantes após o envio dos dados (correspondendo à taxa declarada) também são descartados, abrindo caminho para o próximo conjunto de dados de chegada. Um pacote incompleto é um pacote que não foi totalmente recebido dentro do intervalo de tempo não

é descartado, mas mantido até que tenha sido totalmente recebido na porta.

Esse número de burst supõe um fluxo constante de tráfego. No entanto, nas redes do mundo real, os dados não são constantes e seu fluxo é determinado pelos tamanhos das janelas TCP, que incorporam as confirmações TCP na sequência de transmissão. Para levar em consideração os problemas de tamanhos de janela de TCP, é recomendado que o valor de burst seja dobrado. No exemplo acima, o valor sugerido de 13 seria realmente configurado como 26.

Outro aspecto importante é que no intervalo de tempo 0, ou seja, no início do ciclo de vigilância, o token bucket estará repleto de tokens.

Essa política de agregação agora deve ser incorporada ao ACE de QoS. A ACE é onde a especificação é feita para corresponder um conjunto de critérios a um quadro de entrada. Considere o seguinte exemplo. Você deseja aplicar o agregado definido acima para todo o tráfego IP, mas especificamente para o tráfego com origem da sub-rede 10.5.x.x e com destino para a sub-rede 203.100.45.x. O ACE pareceria com o seguinte:

```
Console> (enable) set qos acl ip test-acl trust-dscp aggregate test-flow tcp 10.5.0.0
203.100.45.0
!-- Test-acl editbuffer modified. Issue the commit command to apply changes.
Console> (enable)
```

O comando acima criou um ACE IP (indicado pelo uso do comando set qos acl ip), que agora está associado a um ACL QoS chamado test-acl. Aces subseqüentes criados e associados ao ACL test-acl são incluídos ao final da lista ACE. A entrada ACE tem o fluxo de teste agregado associado. Qualquer fluxo TCP com uma sub-rede de origem 10.5.0.0 e sub-rede de destino 203.100.45.0 terá essa política aplicada a ele.

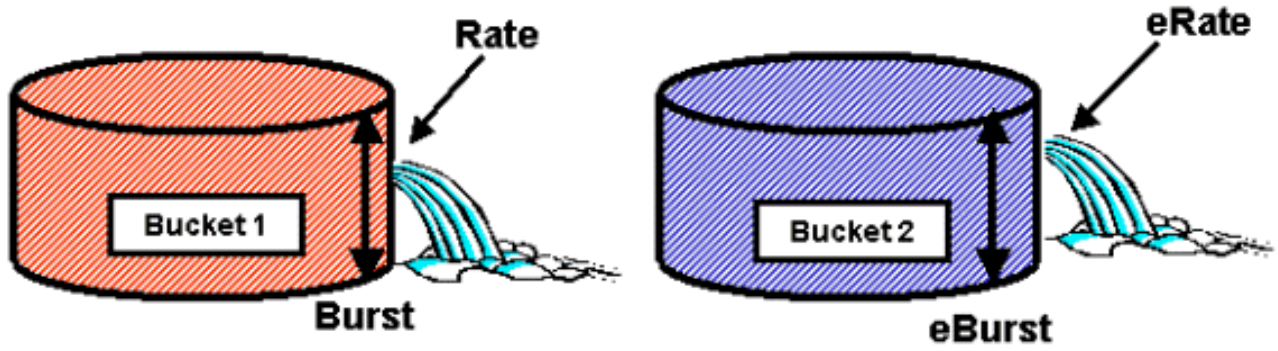
As ACLs (e os Aces associados) fornecem um nível muito granular de flexibilidade de configuração que os administradores podem usar. Uma ACL pode consistir em um ou vários Aces, e os endereços origem e/ou destino podem ser usados, bem como os valores de porta L4 para identificar fluxos específicos que devem ser policiados.

No entanto, antes que qualquer policiamento realmente ocorra, a ACL deve ser mapeada para uma porta física ou para uma VLAN.

Decisões de policiamento de PFC2

Para o PFC2, uma alteração foi feita no CatOS 7.1 e CatOS 7.2, que introduziram um algoritmo de bucket de vazamento duplo para vigilância. Com esse novo algoritmo, ele adiciona os dois novos níveis a seguir:

1. **Nível normal de policiamento:** é igual ao primeiro bucket e define os parâmetros especificando a profundidade do bucket (burst) e a taxa com a qual os dados devem ser enviados do bucket (taxa).
2. **Nível de policiamento de excesso:** isso equivale a um segundo bucket e define parâmetros que especificam a profundidade do bucket (eburst) e a taxa na qual os dados devem ser enviados do bucket (erate).



A forma como esse processo funciona é com os dados começando a preencher o primeiro bucket. O PFC2 aceita um fluxo de entrada de dados menor ou igual à profundidade (valor de intermitência) do primeiro bucket. Os dados que transbordam do primeiro bucket podem ser marcados para baixo e passados para o segundo bucket. O segundo bucket pode aceitar uma taxa de entrada de dados vindos do primeiro bucket em um valor menor ou igual ao valor eburst. Os dados do segundo bucket são enviados a uma taxa definida pelo parâmetro erate menos o parâmetro rate. Os dados que transbordam do segundo bucket também podem ser marcados como inativos ou descartados.

Um exemplo de um vigilante de bucket de dois vazamentos é o seguinte:

```
Console> (enable) set qos policer aggregate AGG1 rate 10000 policed-dscp erate 12000 drop burst
13 eburst 13
```

Esse exemplo posiciona um agregado chamado AGG1 com uma taxa de excesso de tráfego de 10 MPBS e será marcado com um valor inferior de acordo com o mapa de políticas DSCP. O tráfego em excesso do agregado (definido em 12 MBPS) será descartado de acordo com a palavra-chave de perda.

Aplicação de políticas agregadas a módulos habilitados para DFC

Deve-se observar que a aplicação de vigilantes agregados em placas de linha não DFC pode ser alcançada devido à forma como o 6000 usa um mecanismo de encaminhamento centralizado (PFC) para encaminhar tráfego. A implementação de um mecanismo de encaminhamento central permite rastrear as estatísticas de tráfego para uma determinada VLAN. Esse processo pode ser usado para aplicar um vigilante agregado a uma VLAN.

Em uma placa de linha habilitada para DFC, no entanto, as decisões de encaminhamento são distribuídas para essa placa de linha. O DFC só está ciente das portas em sua placa de linha imediata e não está ciente do movimento de tráfego em outras placas de linha. Por esse motivo, se um vigilante agregado for aplicado a uma VLAN que tenha portas membro em vários módulos DFC, o vigilante poderá produzir resultados inconsistentes. A razão para isso é que a DFC só pode controlar as estatísticas de porta local e não leva em conta as estatísticas de porta em outras placas de linha. Por esse motivo, um vigilante agregado aplicado a uma VLAN com portas membro em uma placa de linha habilitada para DFC resultará no tráfego de policiamento de DFC para o limite classificado para portas VLAN residentes somente na placa de linha DFC.

Mapas de marcação DSCP (CatOS)

Os mapas de marcação DSCP são usados quando o vigilante é definido para marcar o tráfego fora de perfil em vez de descartá-lo. O tráfego fora do perfil é definido como o tráfego que excede à configuração definida do surto.

Um mapa padrão de marcação DSCP é configurado quando a QoS está habilitada. Esse mapa

padrão de redução foi listado nessa tabela anteriormente no documento. A CLI (Command Line Interface, interface de linha de comando) permite que um administrador modifique o mapa de marcação padrão emitindo o comando **set qos policed-dscp-map**. Um exemplo é fornecido abaixo.

```
Cat6500(config)# set qos policed-dscp-map 20-25:7 33-38:3
```

Este exemplo modifica o mapa de DSCP policiado para refletir que os valores de DSCP de 20 a 25 serão marcados para um valor de DSCP de 7, e os valores de DSCP de 33 a 38 serão marcados para um valor de DSCP de 3.

Políticas de mapeamento para VLANs e portas (CatOS)

Após a criação de uma ACL, ela deve ser mapeada para uma porta ou uma VLAN para poder ser efetivada.

Um comando interessante que detecta muitos desconhecidos é a configuração padrão de QoS que faz com que toda a porta de QoS seja baseada. Se você aplicar um agregado (ou microfluxo) a uma VLAN, ele não terá efeito em uma porta a menos que essa porta tenha sido configurada para QoS baseada em VLAN.

```
Console> (enable) set port qos 2/5-10 vlan-based
!-- Hardware programming in progress  !-- QoS interface is set to vlan-based for ports 2/5-10.
Console> (enable)
```

A alteração da QoS baseada em porta para a QoS baseada em VLAN desconecta imediatamente todas as ACLs atribuídas a essa porta e atribui qualquer ACL baseada em VLAN a essa porta.

O mapeamento da ACL para uma porta (ou VLAN) é feito emitindo o seguinte comando:

```
Console> (enable) set qos acl map test-acl 3/5
!-- Hardware programming in progress  !-- ACL test-acl is attached to port 3/5. Console>
(enable) Console> (enable) set qos acl map test-acl 18
!-- Hardware programming in progress  !-- ACL test-acl is attached to VLAN 18. Console> (enable)
```

Mesmo depois de mapear a ACL para uma porta (ou uma VLAN), a ACL ainda não entra em vigor até que a ACL seja comprometida com o hardware. Isso está descrito na seção seguinte. Neste ponto, a ACL reside em um buffer de edição temporário na memória. Enquanto estiver nesse buffer, o ACL poderá ser modificado.

Se desejar remover qualquer ACL não comprometida que resida no buffer de editais, você emitirá o comando **rollback**. Esse comando exclui o ACL do buffer de edição.

```
Console> (enable) rollback qos acl test-acl
!-- Rollback for QoS ACL test-acl is successful. Console> (enable)
```

Confirmando ACLs (CatOS)

Para aplicar o QoS ACL que você definiu (acima), o ACL deve estar comprometido com o hardware. O processo de confirmação copia a ACL do buffer temporário para o hardware PFC. Uma vez residindo na memória da PFC, a política definida na ACL de QoS pode ser aplicada a todo o tráfego que corresponda aos ACEs.

Para facilitar a configuração, a maioria dos administradores emite um comando **commit all**. No entanto, você pode confirmar uma ACL específica (uma de muitas) que atualmente pode residir no buffer de edição. Um exemplo do comando commit é mostrado abaixo.

```
Console> (enable) commit qos acl test-acl
!-- Hardware programming in progress  !-- ACL test-acl is committed to hardware. Console>
(enable)
```

Se desejar remover uma ACL de uma porta (ou de uma VLAN), você precisará limpar o mapa que associa essa ACL a essa porta (ou VLAN) emitindo o seguinte comando:

```
Console> (enable) clear qos acl map test-acl 3/5
!-- Hardware programming in progress  !-- ACL test-acl is detached from port 3/5.
Console>(enable)
```

Configure Policing on the Catalyst 6000 Family with Integrated Cisco IOS (Native Mode)

A vigilância é suportada com o Cisco IOS integrado (modo nativo). No entanto, a configuração e a implementação da função de policiamento são obtidas usando mapas de políticas. Cada mapa de política usa várias classes de política para formar um mapa de política e essas classes de política podem ser definidas para diferentes tipos de fluxos de tráfego.

Classes de mapas de política, ao filtrarem, usam ACLs com base em IOS e instruções de correspondência de classe para identificar o tráfego a ser vigiado. Depois que o tráfego for identificado, as classes de política poderão usar os vigilantes agregados de microfluxo para aplicar as políticas de vigilância àquele tráfego correspondente.

As seções seguintes explicam a configuração de vigilância para o Integrated Cisco IOS (Modo Nativo) em mais detalhes.

Agregados e microfluxos (Cisco IOS integrado (modo nativo))

Agregados e microfluxos são termos usados para definir o escopo da vigilância que a PFC executa. Igualmente ao CatOS, os agregados e os microfluxos também são usados como Integrated Cisco IOS (Modo nativo).

Um microfluxo define a vigilância de um único fluxo. Um fluxo é definido por uma sessão com um endereço MAC SA/DA exclusivo, endereço IP SA/DA e números de porta TCP/UDP. Para cada novo fluxo iniciado por uma porta de uma VLAN, o microfluxo pode ser usado para limitar a quantidade de dados recebidos para esse fluxo pelo switch. Na definição de microfluxo, os pacotes que excedem o limite de taxa prescrito podem ser descartados ou ter seu valor de DSCP marcado como inativo. Os microfluxos são aplicados usando o comando police flow que faz parte de uma classe de mapa de política.

To enable microflow policing in Integrated Cisco IOS (Native Mode), it must be enabled globally on the Switch. Isso pode ser feito com a emissão do seguinte comando:

```
Cat6500(config)# mls qos flow-policing
```

O policiamento de microfluxo também pode ser aplicado ao tráfego de ponte, que é o tráfego que não é comutado por L3. Para habilitar o switch a suportar a vigilância de microfluxo no tráfego de ponte, emita o seguinte comando:

```
Cat6500(config)# mls qos bridged
```

Esse comando também permite a vigilância de microfluxo para tráfego multicast. Se o tráfego multicast precisar ter um vigilante de microfluxo aplicado a ele, esse comando (**mls qos bridged**) deverá ser ativado.

Semelhante a um microfluxo, um agregado pode ser usado para limitar a taxa de tráfego. No entanto, a taxa agregada se aplica a todo o tráfego de entrada em uma porta ou VLAN que corresponda a uma ACL de QoS especificada. É possível exibir o agregado como a vigilância de tráfego cumulativo que corresponde a um perfil de tráfego definido.

Existem duas formas de agregados que podem ser definidas no Cisco IOS integrado (modo nativo), como se segue:

- por vigilantes agregados de interface
- vigilantes agregados nomeados

Os agregados por interface são aplicados a um interface individual através da emissão do comando police em uma classe de mapa de política. Essas classes de mapa podem ser aplicadas a várias interfaces, mas o vigilante policia cada interface separadamente. Os agregados nomeados são aplicados a um grupo de portas e policiam o tráfego em todas as interfaces cumulativamente. Os agregados nomeados são aplicados emitindo o comando **mls qos aggregate policer**.

Ao definir microfluxos, pode-se definir até 63 deles e até 1023 agregados.

Criando regras de policiamento (Cisco IOS integrado (modo nativo))

O processo de criação de uma regra de policiamento envolve a criação de um agregado (ou microfluxo) através de um mapa de política e, em seguida, a anexação desse mapa de política a uma interface.

Considere o mesmo exemplo criado para o CatOS. O requisito era limitar todo o tráfego IP de entrada na porta 5/3 a um máximo de 20 MBPS.

Primeiro, um mapa de política deve ser criado. Crie um mapa de política chamado limit-traffic. Isso é feito da seguinte forma:

```
Cat6500(config)# policy-map limit-traffic
Cat6500(config-pmap)#
```

Você observará imediatamente que o prompt do switch muda para refletir que você está no modo de configuração para criar uma classe de mapa. Lembre-se de que um mapa de políticas pode conter múltiplas classes. Cada classe contém um conjunto separado de ações de política que podem ser aplicadas a diferentes fluxos de tráfego.

Devemos criar uma classe de tráfego para limitar especificamente o tráfego recebido a 20 MBPS. Chamaremos essa classe de limite para 20. Isso é mostrado abaixo.

```
Cat6500(config)# policy-map limit-traffic
Cat6500(config-pmap)# class limit-to-20
Cat6500(config-pmap-c)#
```

O prompt se altera novamente para refletir que agora você está na configuração de classe de mapa (mostrado com o -c no fim do prompt). Se quiser aplicar o limite de taxa para corresponder ao tráfego de entrada específico, você pode configurar uma ACL e aplicá-la ao nome da classe. Se quiser aplicar o limite de 20 MBPS ao tráfego originado da rede 10.10.1.x, emita a seguinte ACL:

```
Cat6500(config)# access-list 101 permit ip 10.10.1.0  0.0.0.255 any
```

Você pode adicionar esta ACL ao nome da classe da seguinte maneira:

```
Cat6500(config)# policy-map limit-traffic
Cat6500(config-pmap)# class limit-to-20 access-group 101
Cat6500(config-pmap-c)#
```

Depois que o mapa da classe estiver definido, pode-se definir os vigilantes individuais para essa classe. Você pode criar agregados (usando a palavra-chave "vigia") ou microfluxos (usando a palavra-chave "fluxo de vigias"). Crie o agregado, conforme mostrado abaixo.

```
Cat6500(config)# policy-map limit-traffic
Cat6500(config-pmap)# class limit-to-20 access-group 101
Cat6500(config-pmap-c)# police 20000000 13000 confirm-action transmit exceed-action drop
Cat6500(config-pmap-c)# exit
Cat6500(config-pmap)# exit
Cat6500(config)#
```

A instrução de classe acima (comando police) configura um limite de taxa de 20.000 k (20 Mbps) com um burst de 52 Mbps (13.000 x 4.000 = 52 MB). Se o tráfego corresponder ao perfil e estiver dentro do limite nominal, a ação será definida pela instrução confirm-action para transmitir o tráfego no perfil. Se o tráfego estiver fora de perfil (isto é, em nosso exemplo acima do limite de 20 MB), a instrução de ação excedida será definida para descartar o tráfego (isto é, em nosso exemplo, todo o tráfego acima de 20 MB é descartado).

Na configuração de um microfluxo, uma ação semelhante é executada. Se quiséssemos limitar a taxa de todos os fluxos em uma porta que correspondesse a um determinado mapa de classe a 200 K cada, a configuração desse fluxo seria semelhante à seguinte:

```
Cat6500(config)# mls qos flow-policing
Cat6500(config)# policy-map limit-each-flow
Cat6500(config-pmap)# class limit-to-200
Cat6500(config-pmap-c)# police flow 200000 13000 confirm-action transmit exceed-action drop
Cat6500(config-pmap-c)# exit
Cat6500(config-pmap)# exit
```

Mapas de marcação DSCP

Os mapas de marcação DSCP são usados quando o vigilante é definido para marcar o tráfego fora de perfil em vez de descartá-lo. O tráfego fora do perfil é definido como o tráfego que excede

à configuração definida do surto.

Um mapa de marcação padrão de DSCP é estabelecido quando a QoS está habilitada. Esse mapa de promoção padrão é listado nessa tabela. O CLI permite que um administrador modifique o mapa padrão de redução emitindo o comando set qos policed-dscp-map. Um exemplo é fornecido abaixo.

```
Cat6500(config)#

mls qos map policed-dscp normal-burst 32 to 16
```

Este exemplo define uma modificação no mapa DSCP policed padrão que o valor DSCP de 32 será marcado para um valor DSCP de 16. Para uma porta com esse vigilante definido, qualquer dado recebido com esse valor de DSCP que faça parte de um bloco de dados além da intermitência declarada terá seu valor de DSCP marcado para 16.

Políticas de mapeamento para VLANs e portas (Cisco IOS integrado (modo nativo))

Depois que uma política for criada, ela deverá ser mapeada para uma porta ou para uma VLAN para que ela entre em vigor. Ao contrário do processo de confirmação no CatOS, não há equivalente no Integrated Cisco IOS (modo nativo). Quando uma política é mapeada para uma interface, essa política está em efeito. Para mapear a política acima para uma interface, emita o seguinte comando:

```
Cat6500(config)# interface fastethernet 3/5
Cat6500(config-if)# service-policy input limit-traffic
```

Se uma política for mapeada para uma VLAN, para cada porta na VLAN à qual você deseja que a política de VLAN seja aplicada, você deverá informar à interface que QoS é VLAN com base na emissão do comando **mls qos vlan-based**.

```
Cat6500(config)# interface fastethernet 3/5
Cat6500(config-if)# mls qos vlan-based
Cat6500(config-if)# exit
Cat6500(config)# interface vlan 100
Cat6500(config-if)# service-policy input limit-traffic
```

Supondo que a interface 3/5 fosse parte da VLAN 100, a política denominada limit-traffic que foi aplicada à VLAN 100 também se aplicaria à interface 3/5.

Configure a classificação na família Catalyst 6000 com CatOS

A PFC introduz o suporte para a classificação de dados usando ACLs que podem visualizar informações de cabeçalhos de L2, L3 e L4. Para um SupI, ou IA (sem PFC), a classificação é limitada ao uso das palavras-chave trust nas portas.

A seção a seguir descreve os componentes de configuração QoS usados pelo PFC para Classificação no CatOS.

COs para mapeamento de DSCP (CatOS)

No ingresso no computador, um quadro terá um valor de DSCP definido pelo Switch. Se a porta estiver em um estado confiável e o administrador tiver usado a palavra-chave trust-COs, o valor de COs definido no quadro será usado para determinar o valor de DSCP definido para o quadro. Como mencionado anteriormente, o switch pode atribuir níveis de serviço ao quadro à medida que ele transita o switch com base no valor interno de DSCP.

Esta palavra-chave em alguns dos módulos 10/100 anteriores (WS-X6248 e WS-X6348) não é suportada. Para esses módulos, é recomendável usar ACLs para aplicar configurações de COs para dados de entrada.

Quando o QoS está ativado, o Switch cria um mapa padrão. Esse mapa é usado para identificar o valor DSCP que será definido com base no valor de COs. Esses mapas estão listados nesta tabela anteriormente no documento. Como alternativa, o administrador pode configurar um mapa exclusivo. Um exemplo é fornecido abaixo.

```
Console> (enable) set qos cos-dscp-map 20 30 1 43 63 12 13 8
!-- QoS cos-dscp-map set successfully. Console> (enable)
```

O comando acima configura o seguinte mapa:

CoS	0	1	2	3	4	5	6	7
DSCP	20	30	1	43	63	12	13	8

Enquanto é muito improvável que o mapa acima seja utilizado em uma rede real, ele serve para dar uma idéia do que pode ser alcançado utilizando este comando.

Precedência de IP para mapeamento de DSCP (CatOS)

Da mesma forma que os COs para mapeamento DSCP, um quadro pode ter um valor DSCP determinado a partir da definição de precedência de IP de pacotes recebidos. Isso ainda ocorre somente se a porta estiver definida como confiável pelo administrador e se ele tiver usado a palavra-chave trust-ipprec.

Quando o QoS está ativado, o Switch cria um mapa padrão. Este mapa é citado nessa tabela anteriormente neste documento. Esse mapa é usado para identificar o valor de DSCP que será definido com base no valor de precedência IP. Como alternativa, o administrador pode configurar um mapa exclusivo. Um exemplo é fornecido abaixo:

```
Console> (enable) set qos ipprec-dscp-map 20 30 1 43 63 12 13 8
!-- QoS ipprec-dscp-map set successfully. Console> (enable)
```

O comando acima configura o seguinte mapa:

Precedência de IP	0	1	2	3	4	5	6	7
DSCP	20	30	1	43	63	12	13	8

Enquanto é muito improvável que o mapa acima seja utilizado em uma rede real, ele serve para dar uma idéia do que pode ser alcançado utilizando este comando.

Classificação (CatOS)

Quando um quadro é passado para o PFC para processamento, o processo de classificação é realizado no quadro. O PFC utilizará um ACL pré-configurado (ou um ACL padrão) para atribuir

um DSCP ao quadro. No ACE, uma das quatro palavras-chaves é usada para atribuir um valor de DSCP. São os seguintes:

1. TRUST-DSCP (somente ACLs IP)
2. TRUST-IPPREC (IP ACL'=s only)
3. TRUST-COS (todos os ACLs, exceto IPX e MAC em uma PFC2)
4. DSCP

A palavra-chave TRUST-DSCP pressupõe que o quadro que chega ao PFC já tem um valor de DSCP definido antes de entrar no switch. O Switch manterá esse valor de DSCP.

Com TRUST-IPPREC, a PFC derivará um valor de DSCP do valor de precedência IP existente residente no campo ToS. O PFC utilizará a precedência IP para mapas de DSCP para atribuir o DSCP correto. Um mapa padrão é criado quando a QoS está habilitada no switch. Como alternativa, um mapa criado pelo administrador pode ser usado para derivar o valor de DSCP.

Similarmente à TRUST-IPPREC, a palavra-chave TRUS-COS avisa o PFC para derivar um valor DSCP a partir dos COs no cabeçalho do quadro. Haverá também COs para mapa DSCP (um padrão um de um administrador atribuído a um) para ajudar o PFC na derivação do DSCP.

A palavra-chave DSCP é usada quando um quadro chega a partir de uma porta não-confiável. Isso é uma situação interessante para a derivação do DSCP. Neste ponto, o DSCP configurado na instrução set qos acl é usado para derivar o DSCP. No entanto, é nesse ponto que as ACLs podem ser usadas para derivar um DSCP para tráfego com base nos critérios de classificação definidos na ACE. Isso significa que em um ACE, pode-se usar os critérios de classificação como endereço IP de origem e de destino, números de portas TCP/UDP, códigos ICMP, tipo de IGMP, números de rede e de protocolo IPX, endereços MAC de origem e de destino e Ethertipos (somente para tráfego não-IP e não-IPX) para identificar o tráfego. Isso significa que uma ACE pode ser configurada para atribuir um valor de DSCP específico para dizer tráfego HTTP sobre tráfego FTP.

Considere o seguinte exemplo:

```
Console> (enable) set port qos 3/5 trust untrusted
```

Definir uma porta como não confiável instruirá o PCF a usar um ACE para derivar o DSCP do quadro. Se a ACE for configurada com critérios de classificação, os fluxos individuais dessa porta podem ser classificados com prioridades diferentes. Os Aces a seguir ilustram isso:

```
Console> (enable) set qos acl ip abc dscp 32 tcp any any eq http
Console> (enable) set qos acl ip ABC dscp 16 tcp any any eq ftp
```

Neste exemplo, temos duas afirmações ACE. O primeiro identifica qualquer fluxo TCP (a palavra-chave any é usada para identificar o tráfego de origem e de destino) cujo número de porta é 80 (80 = HTTP) para receber um valor de DSCP de 32. A segunda ACE identifica o tráfego originado de qualquer host e destinado a qualquer host cujo número de porta TCP é 21 (FTP), a quem será atribuído um valor de DSCP de 16.

Configure a classificação da família Catalyst 6000 com Cisco IOS integrado (Modo nativo)

A seção a seguir descreve os componentes de configuração de QoS usados para suportar classificação na PFC usando o Cisco IOS integrado (modo nativo).

COs para mapeamento DSCP (Integrated Cisco IOS (modo nativo))

No ingresso no computador, um quadro terá um valor de DSCP definido pelo Switch. Se a porta estiver em um estado confiável e o administrador tiver usado a palavra-chave mls qos trust-COs (em portas GE ou portas 10/100 nas placas de linha WS-X6548), o valor de COs definido no quadro será usado para determinar o valor de DSCP definido para o quadro. Como mencionado anteriormente, o switch pode atribuir níveis de serviço ao quadro à medida que ele transita o switch com base no valor interno de DSCP.

Quando o QoS está ativado, o Switch cria um mapa padrão. Consulte esta tabela para obter as configurações padrão. Esse mapa é usado para identificar o valor DSCP que será definido com base no valor de COs. Como alternativa, o administrador pode configurar um mapa exclusivo. Um exemplo é fornecido abaixo.

```
Cat6500(config)# mls qos map cos-dscp 20 30 1 43 63 12 13 8
Cat6500(config)#
```

O comando acima configura o seguinte mapa:

CoS	0	1	2	3	4	5	6	7
DSCP	20	30	1	43	63	12	13	8

Enquanto é muito improvável que o mapa acima seja utilizado em uma rede real, ele serve para dar uma idéia do que pode ser alcançado utilizando este comando.

Precedência de IP para mapeamento de DSCP (Cisco IOS integrado (modo nativo))

Da mesma forma que os COs para mapeamento DSCP, um quadro pode ter um valor DSCP determinado a partir da definição de precedência de IP de pacotes recebidos. Isso ainda ocorrerá somente se a porta for definida como confiável pelo administrador e eles tiverem usado a palavra-chave mls qos trust-ipprec. A palavra-chave é suportada apenas em portas GE e 10/100 em placas de linha WS-X6548. Para portas 10/100 das placas de linha WS-X6348 e WS-X6248, as ACLs devem ser usadas para atribuir a confiança de precedência de ip aos dados de entrada.

Quando o QoS está ativado, o Switch cria um mapa padrão. Consulte esta tabela para obter as configurações padrão. Esse mapa é usado para identificar o valor de DSCP que será definido com base no valor de precedência IP. Como alternativa, o administrador pode configurar um mapa exclusivo. Um exemplo é fornecido abaixo.

```
Cat6500(config)# mls qos map ip-prec-dscp 20 30 1 43 63 12 13 8
Cat6500(config)#
```

O comando acima configura o seguinte mapa:

Precedência de IP	0	1	2	3	4	5	6	7
DSCP	20	30	1	43	63	12	13	8

Enquanto é muito improvável que o mapa acima seja utilizado em uma rede real, ele serve para

dar uma idéia do que pode ser alcançado utilizando este comando.

Classificação (Cisco IOS integrado (modo nativo))

Quando um quadro é passado à PFC, o processo de classificação pode ser executado para atribuir uma nova prioridade ao quadro recebido. A advertência nesse caso é que essa atribuição só poderá ser feita quando o quadro for originário de uma porta não confiável ou quando tiver sido classificado como não confiável.

Uma ação de classe de mapa de política pode ser usada para:

1. trust cos
2. TRUST IP-PRECEDENCE
3. TRUST DSCP
4. SEM CONFIANÇA

A palavra-chave de TRUST DSCP supõe que o quadro recebido pelo PFC já possui um valor de DSCP configurado antes de entrar no Switch. O Switch manterá esse valor de DSCP.

Com TRUST IP-PRECEDENCE, o PFC derivará um valor DSCP do valor de precedência de IP existente residente no campo ToS. A PFC usará uma precedência de IP para o mapa DSCP para atribuir o DSCP correto. Um mapa padrão é criado quando a QoS está habilitada no switch. Como alternativa, um mapa criado pelo administrador pode ser usado para derivar o valor de DSCP.

Semelhante à TRUST IP-PRECEDENCE, a palavra-chave TRUST COs instrui o PFC a derivar um valor DSCP dos COs no cabeçalho do quadro. Haverá também COs para mapa DSCP (um padrão um de um administrador atribuído a um) para ajudar o PFC na derivação do DSCP.

Um exemplo de derivação de DSCP de uma prioridade existente (DSCP, precedência de IP ou COs) é mostrado abaixo.

```
Cat6500(config)# policy-map assign-dscp-value
Cat6500(config-pmap)# class test
Cat6500(config-pmap-c)# trust COs
Cat6500(config-pmap-c)# exit
Cat6500(config-pmap)# exit
Cat6500(config)#
```

O mapa de classe acima deduzirá o valor de DSCP dos COs no cabeçalho de Ethernet.

A forma NO TRUST da palavra-chave é usada quando um quadro chega de uma porta não confiável. Isso permite que o quadro tenha um valor de DSCP atribuído durante o processo de policiamento.

Considere o exemplo a seguir de como uma nova prioridade (DSCP) pode ser atribuída a diferentes fluxos que entram no PFC usando a seguinte definição de política.

```
Cat6500(config)# access-list 102 permit tcp any any eq http
Cat6500(config)# policy-map new-dscp-for-flow
Cat6500(config-pmap)# class test access-group 102
Cat6500(config-pmap-c)# no trust
Cat6500(config-pmap-c)# police 1000 1 confirm-action set-dscp-transmit 24  Cat6500(config-pmap-
c)# exit
```

```
Cat6500(config-pmap)# exit
Cat6500(config)#
```

O exemplo acima mostra o seguinte:

1. Uma ACL sendo criada para identificar os fluxos http que entram na porta.
2. Um mapa de política chamado new-dscp-for-flow.
3. Um mapa de classe (teste de nomes) que usa a lista de acesso 102 para identificar o tráfego para o qual esse mapa de classe executará sua ação.
4. O teste de mapa de classe define o estado confiável do quadro de entrada como não-confiável e atribui um DSCP de 24 para o fluxo.
5. Esse mapa de classe também limitará o agregado de todos os fluxos http a um máximo de 1MB.

COPS (Common Open Policy Server)

O COPS é um protocolo que permite que a família Catalyst 6000 tenha a QoS configurada de um host remoto. Atualmente, o COPS só é suportado usando CatOS e faz parte da arquitetura do intserv para QoS. Atualmente, não há suporte (a partir da data deste documento) para COPS ao usar o Cisco IOS Integrado (Modo Nativo). While the COPS protocol carries the QoS configuration information to the Switch, it is not the source of the QoS configuration information. O uso do protocolo COPS requer um gerenciador de QoS externo para hospedar as configurações de QoS do switch. The external QoS manager will initiate the downward push of those configurations to the Switch using the COPS protocol. O QoS Policy Manager (QPM) da Cisco é um exemplo de um QoS Manager externo.

Não é intenção deste documento explicar o funcionamento do QPM, mas explicar a configuração necessária no switch para suportar configurações de QoS externas a partir do uso do QPM.

Configuração do COPS:

Por padrão, o suporte de COPS está desabilitado. Para usar o COPS no switch, ele deve estar ativado. Isso pode ser feito com a emissão do seguinte comando:

```
Console> (enable) set qos policy-source cops
!-- QoS policy source for the switch set to COPS. Console> (enable)
```

Quando este comando é iniciado, determinados valores de configuração de QoS padrão serão originados no servidor COPS. Esses valores de configuração incluem:

1. COs para mapeamentos de fila
2. Atribuições de limites de fila de entrada e saída
3. Atribuições de largura de banda de WRR
4. Qualquer política de agregação e microfluxo
5. Mapas de DSCP para COs para tráfego de saída
6. ACLs
7. Atribuições de COs de porta padrão

Quando as configurações de QoS forem realizadas usando COPS, é importante compreender que o aplicativo dessas configurações é aplicado de uma forma diferente. Mais do que para configurar diretamente as portas, o COPS é utilizado para configurar a porta ASIC. Geralmente, a porta ASIC controla um grupo de portas; portanto a configuração de COPS é aplicada a várias

portas ao mesmo tempo.

A porta ASIC configurada é GE ASIC. Em placas de linha GE, há quatro portas por GE (portas 1-4, 5-8, 9-12, 13-16). Nessas placas de ingresso, a configuração COPS afeta cada grupo de portas. Nas placas de linha 10/100 (conforme discutido anteriormente neste documento), há dois grupos de ASICs, o GE e os ASICs 10/100. Existe um ASIC GE para quatro ASICs 10/100. Cada ASIC 10/100 suporta 12 portas 10/100. O COPS configura o GE ASIC. Assim, ao aplicar a configuração de QoS às placas de linha 10/100 via COPS, a configuração se aplica a todas as 48 portas 10/100.

Ao habilitar o suporte COPS por meio da emissão do comando set qos policy-source cops, a configuração de QoS via COPS será aplicada em todos os ASICs do chassis do Switch. É possível aplicar a configuração de COPS a ASICs específicos. Isso pode ser feito usando o seguinte comando:

```
Console> (enable) set port qos 5/4 policy-source cops
!-- QoS policy source set to COPS for port (s) 5/1-4. Console> (enable)
```

No aplicativo, você pode ver que o comando acima foi emitido em um módulo GE, uma vez que quatro portas foram afetadas por ele.

Servidores de ponto de decisão de política e nomes de domínio

Policy Decision Point Servers (PDPS) are the external policy managers used to store QoS configuration details that are pushed down to the Switch. Se o COPS estiver ativado no switch, o switch deverá ser configurado com o endereço IP do gerenciador externo que fornecerá detalhes de configuração de QoS ao switch. É semelhante a quando o SNMP está ativado e o endereço IP do gerenciador de SNMP está definido.

O comando para identificar o PDPS externo é feito com o uso de:

```
Console> (enable) set cops server 192.168.1.1 primary
!-- 192.168.1.1 is added to the COPS diff-serv server table as primary server. !-- 192.168.1.1
is added to the COPS rsvp server table as primary server. Console> (enable)
```

O comando acima identifica o dispositivo 192.168.1.1 como servidor de ponto de decisão principal.

Quando o switch se comunica com o PDPS, ele precisa fazer parte de um domínio definido no PDPS. O PDPS só falará com os Switches que formam parte de seu domínio definido, portanto o Switch deverá estar configurado para identificar o domínio COPS ao qual ele pertence. Isso é feito pela emissão do seguinte comando:

```
Console> (enable) set cops domain name remote-cat6k
!-- Domain name set to remote-cat6k. Console> (enable)
```

The above command shows the Switch as being configured to be part of the domain named remote-cat6k. This domain should be defined in QPM and the Switch should be added to that domain.

Informações Relacionadas

- Suporte ao Produto - Switches
- Suporte de tecnologia de switching de LAN
- Suporte Técnico e Documentação - Cisco Systems