

Compreendendo limites de fila e quedas de emissor em Plataformas de Cisco IOS Software

Índice

[Introdução](#)

[Pré-requisitos](#)

[Requisitos](#)

[Componentes Utilizados](#)

[Convenções](#)

[Primeira demão do Class-Based Weighted Fair Queueing](#)

[Compreendendo limites de fila e quedas de emissor](#)

[Classes definidas pelo utilizador configuradas com o comando da “prioridade”](#)

[Classes definidas pelo utilizador configuradas com o comando da “largura de banda”](#)

[Comportamento padrão da classe](#)

[Informações Relacionadas](#)

[Introdução](#)

Este documento é aplicável às plataformas de software somente, que inclui geralmente a 3800 do Cisco 7200VXR e do Cisco ISR, 2800 de Cisco IOS®, aos 1800 Series Router, assim como aos roteadores de acesso de Cisco do legado que incluem 3700, 3600, 2600, e 1700 Series Router.

[Pré-requisitos](#)

[Requisitos](#)

A Cisco recomenda que você tenha conhecimento destes tópicos:

- [Soluções da Qualidade de serviço Cisco IOS](#)
- [qos---Estrutura de enfileiramento hierárquica \(HQF\)](#)

[Componentes Utilizados](#)

As informações neste documento são baseadas nestas versões de software:

- Para o PRE-HQF: Roteadores Cisco que executam o Cisco IOS Software Release 12.3(26)
- Para o HQF: Roteadores Cisco que executam o Cisco IOS Software Release 12.4(22)T

As informações neste documento foram criadas a partir de dispositivos em um ambiente de laboratório específico. Todos os dispositivos utilizados neste documento foram iniciados com uma configuração (padrão) inicial. Se a sua rede estiver ativa, certifique-se de que entende o impacto potencial de qualquer comando.

Convenções

Consulte as [Convenções de Dicas Técnicas da Cisco](#) para obter mais informações sobre convenções de documentos.

Primeira demão do Class-Based Weighted Fair Queueing

Nas imagens IOS PRE-HQF, em linhas gerais, qualquer classe com um **comando bandwidth** será dada a prioridade geralmente contra classes sem *largura de banda* ou *prioridade* baseada no *peso das classes*. A fim compreender o algoritmo de escalonamento do Class-Based Weighted Fair Queueing (CBWFQ), você deve primeiramente compreender o conceito de um peso, que seja fluxo-específico para o específico baseado fluxo das filas consideráveis e da classe para classe individual filas baseadas dentro da fila considerável tornada mais pesada baseada da classe.

A fórmula para derivar o peso para uma fila considerável baseada fluxo é:

$32384 / (\text{IP Prec} + 1)$

As classes definidas pelo utilizador dentro de uma fila do equilíbrio justo baseado em classes derivam seus pesos respectivos em função do **comando bandwidth** configurado na classe como uma proporção da SOMA de todas as *classes de largura de banda na fila baseada classe*. A fórmula exata é proprietária.

Em imagens HQF, as filas consideráveis, configuráveis com base no fluxo em ambas as classes e padrão definidos pelo utilizador da classe com *justo-fila*, são programadas igualmente (em vez de por peso). Além disso, no HQF, a prioridade de agendamento de uma fila baseada classe é determinada com base no planificador HQF e não na fórmula do peso do legado das classes.

Nota: Esta seção não é pretendida ser uma análise comportável detalhada para operações de fila baseadas classe. A intenção é uma breve educação porque o CBWFQ se aplica aos limites de fila e às quedas de emissor compreensivos.

Compreendendo limites de fila e quedas de emissor

Classes definidas pelo utilizador configuradas com o comando da “prioridade”

Para as classes definidas pelo utilizador MQC configuradas com alguma variação do **comando priority**, incluindo a **prioridade**, o *<kbps> da prioridade*, e os *por cento da prioridade*.

Comportamento PRE-HQF

Tecnicamente, mesmo que nenhum CLI exista para o ver, e ele não é configurável, “hidden” uma fila de sistema existe que seja compartilhada por todos os dados da classe de prioridade. Isto atua como uma área de terra arrendada para dados da prioridade depois que foi classificado e depois que esteve permitido pelo vigilante congestionamento-ciente. Os pacotes LLQ estão colocados nesta “hidden” fila de sistema se não podem ser colocados diretamente no transmitir anel da interface de saída durante a interrupção da recepção, que está chamada de outra maneira congestionamento funcional. Nesta situação, porque a congestionamento funcional existe, o pacote é avaliado contra o vigilante condicional LLQ durante a interrupção da recepção quando ainda possuído pelo direcionador da relação de recepção. Se o pacote não é deixado cair pelo vigilante condicional do

LLQ, está colocado nesta “hidden” fila de sistema LLQ e a interrupção da recepção é liberada. Consequentemente, todos os pacotes colocados nesta “hidden” fila de sistema conformaram-se ao vigilante ciente da congestão do LLQ e de queueado ao transmitir anel da interface de saída imediatamente pelo planificador LLQ/CBWFQ.

Apesar da existência desta fila, o comportamento é filas desiguais IO criadas para os dados NON-LLQ (tais como a just-to-fila e as filas de largura de banda) que a latência de enfileiramento não adicional (acima da latência do anel de transmissão) estará incorrida desde que os pacotes nesta fila serão drenados imediatamente ao transmitir anel pelo planificador LLQ/CBWFQ. Se um pacote da classe de prioridade não é deixado cair pelo vigilante condicional durante a interrupção da recepção, a seguir este pacote LLQ existirá nesta “hidden” fila de sistema momentaneamente antes de dequeueing ao transmitir anel da interface de saída. Neste caso, o planificador LLQ/CBWFQ apresentará imediatamente o pacote ao transmitir anel da interface de saída. O vigilante condicional foi executado antes de admitir o pacote ao LLQ/CBWFQ, limitando desse modo o LLQ à taxa da prioridade configurada.

Em resumo, recomenda-se deixar cair os dados LLQ que excedem a taxa da prioridade durante épocas da congestão do que para incorrer a latência de enfileiramento adicional, que é um componente fundamental do LLQ. Este mecanismo de policiamento condicional permite uma fila de prioridade estrita sem permitir que a fila de prioridade monopolize a relação inteira PLIM, que é uma melhoria sobre os recursos de enfileiramento de prioridade legado do IOS.

- **Limite de fila PRE-HQF:** NA
- **O PRE-HQF “prioridade aleatório-detecta ” + ”” o comportamento:** NA, WRED não permitido no LLQ.
- **Da “comportamento da just-to-fila prioridade PRE-HQF ” + ””:** NA, just-to-fila não permitida no LLQ.
- **A “prioridade aleatório-detecta ” + ” ” + ” comportamento da just-to-fila PRE-HQF”:** O NA, nem just-to-fila ou aleatório-detecta apoiado no LLQ.

[Comportamento HQF](#)

Mesmo como PRE-HQF a não ser que hidden a fila seja já não hidden e o fila-limite é agora configurável e opta 64 pacotes.

- **Limite de fila HQF:** 64 pacotes
- **O HQF “prioridade aleatório-detecta ” + ”” o comportamento:** NA, WRED não permitido no LLQ.
- **Comportamento da just-to-fila prioridade HQF da “” + ””:** NA, just-to-fila não permitida no LLQ.
- **Comportamento da just-to-fila HQF a “prioridade aleatório-detecta ” + ” ” + ””:** O NA, nem just-to-fila ou aleatório-detecta apoiado no LLQ.

[Classes definidas pelo utilizador configuradas com o comando da “largura de banda”](#)

Para as classes definidas pelo utilizador MQC configuradas com alguma variação do comando **bandwidth**, incluindo o **<kbps>**, o **percentagem de largura de banda**, e o **percentagem restante de largura de banda da largura de banda**.

[Comportamento PRE-HQF](#)

O limite de fila padrão é 64 pacotes, que é ajustável. Se, durante a interrupção da recepção, você precisa de enviar à fila um pacote que conduza > a 64 pacotes na fila, o pacote é cauda deixada cair.

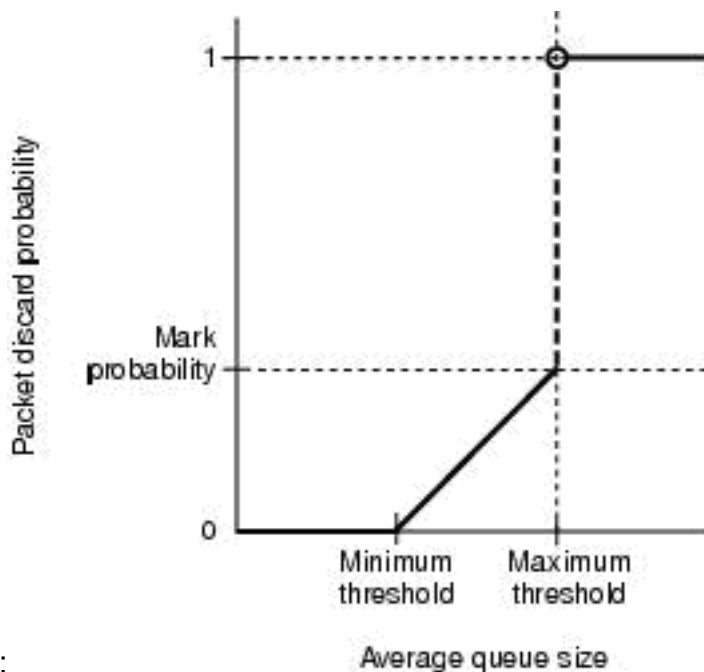
- **Fila-limite PRE-HQF:** 64 pacotes, ajustáveis através do fila-limite.
- **O PRE-HQF "largura de banda aleatório-detecta" + " o comportamento:**Exemplo:policy-map

```
PRE_HQF_BANDWIDTH_WRED
```

```
class 1  
  bandwidth 32
```

```
  random-detect
```

Se qualquer variação da largura de banda é configurada junto com qualquer variação de aleatório-detecte, ignore todo o fila-limite CLI, que remover eficazmente qualquer limite de buffer na classe. Ou seja aleatório-detecte e o fila-limite é mutuamente exclusivos nas imagens PRE-HQF. Usando-se aleatório-detecte como a estratégia da gota, o tamanho da fila atual é desenfreado e pode teoricamente ocupar cada buffer atribuído à justo-fila baseada classe, onde este número de buffer atribuído à justo-fila baseada classe é derivado baseou no ponto de anexo da serviço-política: Interface física: 1000 pacotes, ajustáveis com a posse-fila da relação CLI para fora ATM PVC: 500 pacotes, ajustáveis com a VC-posse-fila PVC CLI Classe de mapas do Frame Relay: 600 pacotes, ajustáveis com holdq do Frame Relay da classe de mapas CLI do Frame Relay Política do modelagem baseada em classe aplicada à relação (secundária) (PRE-HQF): 1000 pacotes, ajustáveis com MAX-buffers da forma da classe CLI MQC. **Nota:** Todo o Frame Relay e classe baseados dando forma a exemplos supõem que a soma dos shapers não excede o Clock Rate da relação. **Nota:** Nas imagens PRE-HQF, quando uma política moldada baseada classe está aplicada à relação (secundária) a, ter cuidado com a velocidade da interface física subjacente, como as relações <2Mbps optarão uma fila considerável tornada mais pesada e as relações >2Mbps optará o FIFO. No pre-HQF, a fila moldada alimentará a fila de contenção da relação, se a política moldada está anexada a subinterface ou nível da interface física. Durante a interrupção da recepção, cada vez que um pacote assenta bem em um candidato para a fila de saída de uma relação, o tamanho médio da fila WRED é calculado usando esta fórmula: $Average\ Queue\ Size = (old_average * (1 - 1/2^n)) + (current_queue_size * 1/2^n)$ Se o tamanho médio da fila resultante é: Menor do que o limiar mín WRED, envie à fila o pacote e libere a interrupção da recepção. Entre o limiar mín WRED e o limiar máx WRED, deixe cair possivelmente o pacote com probabilidade aumentada como o tamanho médio da fila obtém mais perto do limiar máx WRED. Se o tamanho médio da fila é exatamente igual ao limiar máx WRED, o pacote está deixado cair de acordo com o denominador de probabilidade da marca. O denominador de probabilidade da marca igualmente serve como uma linha de base para determinar que porcentagem de pacotes obterá deixado cair quando o limite da fila média não é exatamente igual ao limiar máx WRED mas é mais alto do que o limiar mín WRED. Este é um exemplo



10/7/8

gráfico:

Se o pacote é deixado cair, a interrupção da recepção está liberada e uma gota aleatória é incrementada. Se o pacote não é deixado cair, o pacote está enviado à fila e a interrupção da recepção é liberada. Mais altamente do que o limiar máx WRED, deixe cair o pacote, libere a interrupção da recepção, e incremente uma queda traseira.

- **Nota:** O WRED baseado (padrão) e DSCP-baseado da Precedência IP permite o limiar mín, o limiar máx, e o denominador de probabilidade da marca a ser definidos diferentemente para valores diferentes. Isto é o lugar onde o componente tornado mais pesado da detecção adiada aleatória é evidente. Você pode proteger determinados valores ToS relativo a outros valores através de ajustar seus pontos iniciais relativos e marcar denominadores de probabilidade. Quando aleatório detecte e a largura de banda é configurada junto, o tamanho da fila atual pode ser maior do que o limiar máx WRED em algum ponto dado a tempo. Isto é porque WRED mínimo e limiares máximos tome somente a ação baseada no tamanho da fila (não atual) médio. Isto fornece uma oportunidade de expirar todos os buffers atribuídos à fila baseada classe que poderia conduzir às gotas do sem buffer que ocorrem em qualquer lugar dentro da fila considerável baseada classe (refira a identificação de bug Cisco CSCsm94757).
- Da "comportamento da justo-fila largura de banda PRE-HQF " + """: NA, justo-fila não permitida na classe de largura de banda.
- A "largura de banda aleatório-detecta " + "" + " comportamento da justo-fila PRE-HQF": NA, justo-fila não permitida na classe de largura de banda

Comportamento HQF

O comportamento é o mesmo como descrito na seção PRE-[HQF](#).

- **Fila-limite HQF:** 64 pacotes, ajustáveis através do fila-limite. Este é mesmo que aquele no PRE-HQF.
- O HQF "largura de banda aleatório-detecta " + "" o comportamento: Exemplo: `policy-map`

`HQF_BANDWIDTH_WRED`

```
class 1
  bandwidth 32
  queue-limit 512
  random-detect
```

Nota: O limite de fila padrão é 64 pacotes. O comportamento é o mesmo que na seção PRE-HQF correspondente, com uma exceção importante. Em imagens HQF,

aleatório-detecte e o fila-limite pode coexistir na mesma classe definida pelo utilizador (ou no class class-default) e o fila-limite será permitido e ajustado a 64 pacotes em uma configuração padrão. Como tal, o fila-limite servirá como um tamanho da fila atual máximo em uma classe da aleatório-detecteção, consequentemente fornecendo um mecanismo para limitar as gotas do sem buffer discutidas na seção PRE-HQF correspondente. Devido a esta adição, o limite de fila configurado deve ser pelo menos tão grande quanto o limiar máx da aleatório-detecteção, onde o limiar máx da aleatório-detecteção optará 40 pacotes, ou então o parser rejeitarão a configuração. Isto introduz um atual-fila-limite verifica dentro classes WRED, por meio de que, mesmo se o cálculo da profundidade média da fila é menor do que o limiar máx, se o tamanho da fila (não médio) atual é maior do que o fila-limite, o pacote será deixado cair, a interrupção da recepção serão liberadas, e uma queda traseira gravada. Recorde, se o fila-limite é ajustado suficientemente altamente para esgotar os buffers de enfileiramento do agregado para a fila com base na classe, gotas do sem buffer pode ainda ocorrer. Os buffers de enfileiramento agregados para o HQF são definidos aqui: Interface física: 1000 pacotes, ajustáveis com a posse-fila da relação CLI para fora ATM PVC: 500 pacotes, ajustáveis com a VC-posse-fila PVC CLI Classe de mapas do Frame Relay: 600 pacotes, ajustáveis com holdq do Frame Relay da classe de mapas CLI do Frame Relay Política do modelagem baseada em classe aplicada à interface física no código HQF: 1000 pacotes, ajustáveis com uma combinação de posse-fila da relação CLI para fora e de fila-limite da política infantil onde o fila-limite da política infantil tem um limite superior da posse-fila da relação para fora. Política do modelagem baseada em classe aplicada à subinterface no código HQF: 512 pacotes, não ajustáveis (investigue com a equipe da plataforma NSSTG QoS, se for ajustável) Nota: Todo o Frame Relay e classe baseados dando forma a exemplos supõem que a soma dos shapers não excede o Clock Rate da relação. Está aqui um exemplo do mundo real:

```

policy-map JACKLYN
class 1
  bandwidth 64
  queue-limit 500 packets
  random-detect
  random-detect precedence 1 22 300

```

Durante esta saída, o sem tráfego está sendo gerado através da relação:

```

F340.11.25-7200-5_LAC#show policy-map interface | i queue
queue limit 500 packets
(queue depth/total drops/no-buffer drops) 0/387595/0
!--- Current_q_depth is 0 Mean queue depth: 107 packets !--- last calculation of
Average_queue_depth

```

Neste momento, o tráfego é começado. O córrego é NON-intermitência em 400PPS, os frames de bytes of 1000 consistindo:

```

F340.11.25-7200-5_LAC#show policy-map
interface | i queue
queue limit 500 packets
(queue depth/total drops/no-buffer drops) 461/387641/0
!--- 461 is Current_q_depth > Prec 1 max-thresh of 300 !--- but < "queue-limit 500 packets".
Mean queue depth: 274 packets !--- Avg_q_depth is rising, Mark Prob Denom is being used.
F340.11.25-7200-5_LAC#show policy-map interface | i queue queue limit 500 packets (queue
depth/total drops/no-buffer drops) 363/387919/0 !--- 363 is Current_q_depth and it is
falling compared to last !--- iteration because WRED is random dropping packets. Mean queue
depth: 351 packets !--- Avg_q_depth is now above max_thresh, WRED starts to tail-drop !---
in addition to random-drop. F340.11.25-7200-5_LAC#show policy-map interface | i queue queue
limit 500 packets (queue depth/total drops/no-buffer drops) 199/388263/0 Mean queue depth:
312 packets F340.11.25-7200-5_LAC#show policy-map interface | i queue queue limit 500
packets (queue depth/total drops/no-buffer drops) 303/388339/0 Mean queue depth: 276 packets
F340.11.25-7200-5_LAC#show policy-map interface | i queue queue limit 500 packets (queue
depth/total drops/no-buffer drops) 325/388498/0 Mean queue depth: 314 packets F340.11.25-
7200-5_LAC#show policy-map interface | i queue queue limit 500 packets (queue depth/total
drops/no-buffer drops) 298/390075/0 Mean queue depth: 300 packets

```

Observação como,

eventualmente, com um córrego da NON-intermitência, a profundidade média da fila WRED igualará a profundidade de fila atual, que é o comportamento esperado.

- **Comportamento da justo-fila largura de banda HQF da " " + " ":** Quando a largura de banda e a justo-fila são aplicadas junto a uma classe definida pelo utilizador HQF, cada fila com base no fluxo está atribuída um fila-limite igual a $.25 * \text{fila-limite}$. Porque o limite de fila padrão é 64 pacotes, cada fila baseada fluxo em uma justo-fila será atribuída 16 pacotes. Se quatro fluxos atravessavam esta classe, à revelia cada fila de fluxo teria 16 pacotes, conseqüentemente você nunca esperaria ver os pacotes total enviados à fila de $>64 (4 * 16)$. Todas as quedas traseiras de uma fila de fluxo individual são gravadas como flowdrops. Se o número de filas de fluxo era significativamente alto como era o fila-limite, então uma outra oportunidade para o sem buffer deixa cair. Por exemplo, o anexo-ponto presumido da política é uma interface física, onde 1000 buffers agregados sejam atribuídos:`policy-map TEST`

```
class 1
  bandwidth 32
  fair-queue 1024
```

`queue-limit 128` Nesta configuração, o tráfego apreciável em todas as filas de fluxo pode morrer de fome buffers e o resultado agregados da relação em gotas do sem buffer em outras classes definidas pelo utilizador (veja a identificação de bug Cisco CSCsw98427). Isto é porque 1024 filas de fluxo, cada um com um fila-limite de 32 pacotes podem facilmente oversubscribe a alocação de buffer de enfileiramento baseada da relação 1000 classe agregada.

- **Comportamento da justo-fila HQF a "largura de banda aleatório-detecta " + " " + " ":** Exemplo:`policy-map TEST`

```
class 1
  bandwidth 32
  fair-queue 1024
  queue-limit 128
```

`random-detect` Mesmo que a largura de banda e a justo-fila na seção a não ser que o tamanho médio da fila WRED seja calculado todas as vezes um pacote chega para decidir se o pacote deve ser deixado cair aleatório ou cauda deixada cair. Como com PRE-HQF, todas as filas de fluxo compartilharão de um exemplo dos limites de WRED, significando todos os pacotes enviados à fila a todas as filas de fluxo são usados para calcular a profundidade média da fila WRED, a seguir a decisão da gota aplica o WRED mínimo e os limiares máximos contra os pacotes agregados em todas as filas de fluxo. Contudo, uma outra partida da largura de banda e a justo-fila na seção, porque um exemplo dos limites de WRED se aplica a todas as filas com base no fluxo, o fila-limite das filas de fluxo individuais ($.25 * \text{"fila-limite"}$) são ignoradas e honram pelo contrário o fila-limite agregado das classes para uma verificação de limite de fila atual.

[Comportamento padrão da classe](#)

[Comportamento PRE-HQF](#)

No pre-HQF, o padrão da classe opta pela justo-fila, todas as filas de fluxo compartilham do fila-limite para a classe (o padrão é 64), e não há nenhuma reserva de largura de banda. Ou seja o número total de pacotes enviados à fila em todas as filas de fluxo nunca excederá o fila-limite. A quantidade de pacotes enviados à fila em cada fila de fluxo dependerá do peso calculado da fila de fluxo. Inversamente, se a justo-fila e aleatório-detecta é usada junto no class-default, o fila-limite será ignorado e todas as filas de fluxo compartilharão dos mesmos limites de WRED. Como tal, todos os pacotes enviados à fila atualmente em todas as filas de fluxo serão usados para

calcular o tamanho médio da fila WRED. Porque o tamanho da fila atual não tem nenhum limite superior nesta configuração, a oportunidade para gotas do sem buffer é alta (refira a identificação de bug Cisco CSCsm94757).

- Adicionar a largura de banda ao class-default conduzirá ao comportamento esboçado no [comportamento PRE-HQF - as classes definidas pelo utilizador configuradas com o comando section da "largura de banda"](#).
- Adicionando a largura de banda e aleatório-detecte ao class class-default conduzirá ao comportamento esboçado no PRE-HQF *"largura de banda aleatório-detectam " + ""* a seção do *comportamento do* [comportamento PRE-HQF - as classes definidas pelo utilizador configuradas com o comando da "largura de banda"](#).

Nota: No pre-HQF, a largura de banda não pode coexistir com a justo-fila no class-default.

[Comportamento HQF](#)

No HQF, o padrão da classe opta uma fila de FIFO e é atribuído uma reserva de largura de banda pseudo- baseada nas atribuições restantes das classes definidas pelo utilizador. Como tal, para o comportamento padrão do class class-default HQF, veja o [comportamento HQF - as classes definidas pelo utilizador configuradas com o comando section da "largura de banda"](#). Em todas as vezes, apesar da configuração, o class class-default em imagens HQF terá sempre uma reserva de largura de banda implícita igual à largura de banda da interface não utilizada não consumida por classes definidas pelo utilizador. À revelia, a classe de padrão classe recebe um mínimo de 1% da largura de banda da forma da relação ou do pai. É igualmente possível configurar explicitamente a largura de banda CLI no padrão da classe.

- Se a justo-fila é configurada no class class-default, o comportamento combina a seção do *comportamento da justo-fila largura de banda HQF da "" + ""* do [comportamento HQF - as classes definidas pelo utilizador configuradas com o comando da "largura de banda"](#).
- Se a justo-fila e aleatório-detecta é configurada junto no class-default, o comportamento combina a seção do *comportamento da justo-fila HQF "largura de banda aleatório-detecta " + "" + ""* do [comportamento HQF - as classes definidas pelo utilizador configuradas com o comando da "largura de banda"](#).

[Informações Relacionadas](#)

- [Manual de configuração das soluções da Qualidade de serviço Cisco IOS, liberação 12.4T](#)
- [qos---Estrutura de enfileiramento hierárquica \(HQF\)](#)
- [Suporte da tecnologia de QoS](#)
- [Sustentação do produto do Roteadores](#)
- [Suporte Técnico e Documentação - Cisco Systems](#)