

# Solucionar problemas do roteador na rede corporativa

## Contents

---

[Introdução](#)

[Pré-requisitos](#)

[Requisitos](#)

[Componentes Utilizados](#)

[Informações de Apoio](#)

[Definição de latência](#)

[Uso de latência](#)

[Problemas de latência próximos](#)

[Solucionar causas comuns](#)

[Relacionado à plataforma](#)

[Alta utilização da CPU](#)

[Tráfego relacionado](#)

[MTU e Fragmentação](#)

[Relacionado ao design](#)

[Roteamento não ideal](#)

[Quality of Service \(QoS\)](#)

[Outros problemas de desempenho](#)

[Quedas](#)

[Retransmissão TCP](#)

[Excesso de demanda e gargalos](#)

[Informações Relacionadas](#)

---

## Introdução

Este documento descreve como identificar, solucionar problemas e resolver problemas de latência em redes corporativas usando roteadores Cisco.

## Pré-requisitos

### Requisitos

Não há pré-requisitos ou requisitos específicos para este documento.

### Componentes Utilizados

Este documento não se restringe à versão de software e ao tipo de hardware específicos, mas os comandos são aplicáveis aos roteadores Cisco IOS® XE como as famílias ASR 1000, ISR 4000 e

Catalyst 8000.

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 rede estiver ativa, certifique-se de que você entenda o impacto potencial de qualquer comando.

## Informações de Apoio

Este documento descreve um guia básico para entender, isolar e solucionar problemas gerais de latência, fornece comandos/depurações úteis para detectar as causas raiz e as práticas recomendadas. Tenha em mente que todas as variáveis e cenários possíveis não podem ser considerados e uma análise mais profunda depende de situações específicas.

## Definição de latência

Em termos gerais, e citando a definição estrita para dispositivos de armazenamento e encaminhamento (no RFC 1242), latência é o intervalo de tempo que começa quando o último bit do quadro de entrada alcança a porta de entrada e termina quando o primeiro bit do quadro de saída é visto na porta de saída.

A latência de rede pode simplesmente se referir ao atraso na transferência de dados através da rede. Para questões práticas, essa definição é apenas o ponto de partida; você precisa definir o problema de latência do qual você está falando em cada caso específico, embora pareça óbvio, o primeiro passo necessário para resolver um problema, e se torna realmente importante, é defini-lo.

## Uso de latência

Muitos aplicativos requerem baixa latência para comunicação em tempo real e operações de negócios; com as melhorias de hardware e software todos os dias, mais aplicativos estão disponíveis para computação de missão crítica, aplicativos de reunião on-line, transmissão, entre outros; da mesma forma, o tráfego de rede continua a crescer e a necessidade de projetos de rede otimizados e melhor desempenho de dispositivos também aumenta.

Além de proporcionar uma melhor experiência ao usuário e fornecer o mínimo necessário para aplicativos sensíveis à latência, identificar e reduzir com eficiência problemas de latência em uma rede pode economizar muito tempo e recursos altamente valiosos em uma rede.

## Problemas de latência próximos

A parte difícil desse tipo de problema é o número de variáveis que você deve levar em consideração, além de não haver um único ponto de falha. Portanto, a definição de latência se torna uma chave importante para resolvê-la e alguns aspectos que devem ser considerados para que uma descrição útil do problema seja feita são os próximos.

## 1. Expectativa e detecção

É importante diferenciar uma latência desejada, a latência de trabalho esperada ou de linha de base e a atual. Dependendo do design, dos provedores ou dos dispositivos na rede, às vezes você não pode atingir a latência desejada, é um bom procedimento para medir a latência real em condições normais, mas você precisa ser consistente com os métodos de medição para evitar números enganosos; SLAs IP e ferramentas de analisador de rede podem ajudar nesse aspecto.

Uma das ferramentas mais usadas e básicas para identificar a latência por aplicativos ou mesmo SLA IP é via ICMP ou ping:

```
<#root>
Router#
ping
 198.51.100.1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 198.51.100.1, timeout is 2 seconds:
!!!!
Success rate is 100 percent (5/5),
round-trip min/avg/max
=
2/109/541 ms
```

Além de verificar a acessibilidade, o ping informa o RTT (Round Trip Time, Tempo de ida e volta) da origem para o destino; o mínimo (2), a média (109) e o máximo (541) em milissegundos. Isso significa a duração do envio da solicitação pelo roteador até o recebimento da resposta do destino do dispositivo. No entanto, ele não mostra quantos saltos ou informações mais detalhadas, mas é uma maneira fácil e rápida de detectar um problema.

## 2. Isolamento

Assim como o ping, o traceroute pode ser usado como o ponto de partida para o isolamento, ele descobre os saltos e o RTT por salto:

```
<#root>
Router#
traceroute
 198.51.100.1
Type escape sequence to abort.
Tracing the route to 198.51.100.1
VRF info: (vrf in name/id, vrf out name/id)
 1 10.0.3.1 5 msec 6 msec 1 msec
 2 10.0.1.1 1 msec 1 msec 1 msec
 3 10.60.60.1 1 msec 1 msec 1 msec
```

4 10.90.0.2

362 msec 362 msec 362 msec

<<<< you can see the RTT of the three probes only on both hops

5 10.90.1.2

363 msec 363 msec 183 msec

6 10.90.7.7 3 msec 2 msec 2 msec

O Traceroute opera enviando um pacote com um TimeTo Live (TTL) de 1. O primeiro salto envia de volta uma mensagem de erro ICMP indicando que o pacote não pôde ser encaminhado porque o TTL expirou e o RTT foi medido, o segundo pacote é reenviado com um TTL de 2 e o segundo salto retorna o TTL expirado. Esse processo continua até que o destino seja alcançado.

No exemplo, agora você pode se restringir a dois hosts específicos e pode começar a partir daí em nosso isolamento.

Apesar de serem comandos úteis que podem identificar facilmente um problema, eles não levam em consideração outras variáveis, como protocolos, marcações e tamanhos de pacotes (embora você possa defini-los como segunda etapa), diferentes origens IP, destinos entre vários fatores.

Dizer latência pode ser um conceito muito amplo e você frequentemente vê apenas o sintoma em um aplicativo, navegação, chamada ou tarefas específicas. Uma das primeiras coisas a limitar é entender o impacto e definir o problema com mais detalhes, responder às próximas perguntas e os elementos podem ajudar nesse dimensionamento:

- A latência afeta apenas um tipo específico de tráfego ou aplicativo? Exemplo: somente UDP, TCP, ICMP...
- Em caso afirmativo, esse tráfego tem identificadores exclusivos? Exemplo: marcação de QoS específica, somente tamanhos de pacotes determinados, opções de IP...
- Quantos usuários ou locais são afetados? Exemplo: apenas uma sub-rede específica, um ou dois hosts finais, um site inteiro conectado a um ou vários dispositivos...
- Há carimbos de data/hora específicos identificados? Exemplo: isso ocorre apenas durante as horas de pico, qualquer padrão de tempo ou aleatório completo...
- Aspectos de projeto. Exemplo: o tráfego que passa por um dispositivo específico, talvez muitos dispositivos, mas se conecta a apenas um provedor, o tráfego que faz o balanceamento de carga, mas afetou um caminho...

Há muitas outras considerações, mas cruzar as diferentes respostas (e até mesmo testes que podem ser feitos para respondê-las) pode isolar e limitar efetivamente o escopo para prosseguir com a solução de problemas. Por exemplo, apenas um aplicativo (tráfego do mesmo tipo) foi afetado em todas as filiais que passavam por diferentes provedores e terminavam no mesmo data center no horário de pico. Nesse caso, você não começa a verificar todos os switches de acesso em todas as filiais. Em vez disso, você se concentra em coletar mais informações sobre o data center e inspeciona mais nesse lado,

As ferramentas de monitoramento e alguma automação que você pode ter na rede também ajudam muito nesse isolamento, realmente depende dos recursos que você tem e de situações únicas.

## Solucionar causas comuns

Depois de limitar o escopo da solução de problemas, você pode começar a verificar causas específicas, por exemplo, no exemplo de traceroute fornecido, você pode isolar para dois saltos diferentes e, em seguida, restringir para possíveis causas.

### Relacionado à plataforma

#### Alta utilização da CPU

Uma das causas comuns pode ser um dispositivo com alto atraso de criação de CPU no processo de todos os pacotes. Para roteadores, os comandos mais úteis e básicos para verificar os roteadores são

Desempenho geral do roteador:

```
<#root>
```

```
Router#
```

```
show platform resources
```

```
**State Acronym: H - Healthy, W - Warning, C - Critical
```

Resource	Usage	Max	Warning	Critical	State
-----					
RPO (ok, active)					H
Control Processor	1.15%	100%	80%	90%	H
DRAM	3631MB (23%)	15476MB	88%	93%	H
bootflash	11729MB (46%)	25237MB	88%	93%	H
harddisk	1121MB (0%)	225279MB	88%	93%	H
ESP0(ok, active)					H
QFP					H
TCAM	8cells (0%)	131072cells	65%	85%	H
DRAM	359563KB (1%)	20971520KB	85%	95%	H
IRAM	16597KB (12%)	131072KB	85%	95%	H
CPU Utilization	0.00%	100%	90%	95%	H
Crypto Utilization	0.00%	100%	90%	95%	H

Pkt Buf Mem (0)	1152KB(0%)	164864KB	85%	95%	H
Pkt Buf CB1k (0)	14544KB(1%)	986112KB	85%	95%	H

Útil para ver a utilização da memória e da CPU de uma só vez, ele é dividido no plano de controle e no plano de dados (QFP), iguais aos limites de cada um. A memória em si não cria um problema de latência; no entanto, se não houver mais memória DRAM para o plano de controle, o Cisco Express Forwarding (CEF) será desabilitado e induzirá um alto uso da CPU que pode produzir latência, por isso é importante manter os números em estado íntegro. O guia básico para solução de problemas de memória está fora do escopo, mas consulte o link útil na seção Informações relacionadas.

Se for detectada alta utilização da CPU para o Processador de Controle, CPU QFP ou Criptografia, você poderá usar os seguintes comandos:

Para o plano de comando:

```
show process cpu sorted
```

```
<#root>
```

```
Router#
```

```
show processes cpu sorted
```

```
CPU utilization for five seconds:
```

```
99%/0%
```

```
; one minute: 13%; five minutes: 3%
```

PID	Runtime(ms)	Invoked	uSecs	5Sec	1Min	5Min	TTY	Process
65	1621	638	2540	89.48%	1.82%	0.41%	0	crypto sw pk pro
9	273	61	4475	1.56%	0.25%	0.05%	0	Check heaps
51	212	64	3312	0.72%	0.21%	0.05%	0	Exec
133	128	16	8000	0.60%	0.08%	0.01%	0	DBAL EVENTS
473	25	12	2083	0.48%	0.04%	0.00%	0	WSMAN Process
84	1173	353	3322	0.36%	0.07%	0.02%	0	IOSD ipc task
87	23	12	1916	0.24%	0.02%	0.00%	0	PuntInject Keepa
78	533	341	1563	0.12%	0.29%	0.07%	0	SAMsgThread
225	25	1275	19	0.12%	0.00%	0.00%	0	SSS Feature Time
386	4	4	1000	0.12%	0.00%	0.00%	0	Crypto WUI
127	204	18810	10	0.12%	0.02%	0.00%	0	L2 LISP Punt Pro

Se a CPU do plano de controle estiver alta (este exemplo está em 99% devido a processos), será necessário isolar o processo e, dependendo dele, continuar com o isolamento (podem ser pacotes punted para nós como ARP ou pacotes de rede de controle, podem ser qualquer protocolo de roteamento, multicast, NAT, DNS, tráfego de criptografia ou qualquer serviço).

Dependendo do seu fluxo de tráfego, isso pode causar um problema no processamento posterior. Se o tráfego não for destinado ao roteador, você pode se concentrar no plano de dados:

Para o plano de dados:

show platform hardware qfp active datapath usage [summary]

<#root>

Router#

show platform hardware qfp active datapath utilization

CPP 0: Subdev 0

5 secs

	1 min	5 min	60 min		
Input: Priority	(pps)	0	0	0	0
	(bps)	0	0	0	0
Non-Priority	(pps)	231	192	68	6
	(bps)	114616	95392	33920	3008
Total	(pps)	231	192	68	6
	(bps)	114616	95392	33920	3008
Output: Priority	(pps)	0	0	0	0
	(bps)	0	0	0	0
Non-Priority	(pps)	3	2	2	0
	(bps)	14896	9048	8968	2368

Total (pps)

3323 2352 892 0

(bps)

14896 9048 8968 2368

Processing: Load (pct)

3

3 3 3

Crypto/I0

Crypto: Load (pct)

0

0	0	0	0	0	0
RX: Load (pct)		0	0	0	0
TX: Load (pct)		1	1	0	0
Idle (pct)		99	99	99	99

Se o plano de dados for alto (identificado pelo número da carga de processamento atingindo 100%), será necessário ver a quantidade de tráfego que passa pelo roteador (pacote total por segundos e bits por segundo) e o desempenho de throughput da plataforma (você pode ter uma ideia sobre a folha de dados específica).

Para determinar se esse tráfego é esperado ou não, a captura de pacotes (EPC) ou qualquer recurso de monitoramento, como o Netflow, pode ser usado para análise adicional, algumas verificações são:

- O tráfego é válido e deve passar por esse roteador?
- Identificar fluxos de tráfego anormais ou taxas mais altas.
- Se você tiver números altos de pacotes por segundo, procure o tamanho dos pacotes. Determine se isso é esperado para ver ou se há um problema de fragmentação.

Se todo o tráfego for esperado, você pode estar atingindo uma limitação de plataforma, então, procure os recursos em execução no seu roteador como uma segunda parte para análise através de show running-config, principalmente nas interfaces, identifique quaisquer recursos desnecessários e desative-os ou equilibre o tráfego para liberar ciclos de CPU.

No entanto, se não houver indicação de um limite de plataforma, outra ferramenta útil para confirmar se o roteador está adicionando atraso em pacotes é o rastreamento FIA, você pode ver o tempo de processo exato gasto para cada pacote e os recursos que estão fazendo a maior parte do processamento. A solução de problemas de alta CPU está fora do escopo deste documento, mas consulte os links na seção Informações Relacionadas.

## Tráfego relacionado

### MTU e Fragmentação

Maximum Transmission Unit (MTU) é o comprimento máximo do pacote a ser transmitido, que depende do número de octetos que os links físicos podem transmitir. Quando os protocolos de camada superior enviam dados ao IP subjacente e o comprimento resultante do pacote IP é maior que o MTU do caminho, o pacote é dividido em fragmentos. Esses tamanhos menores na rede causam mais processamento e tratamento diferente em alguns casos e é por isso que você deve evitá-los o mais possível.

Para alguns recursos, como NAT ou Firewall baseado em zona, a remontagem virtual é necessária para "ter o pacote inteiro", aplica o que é necessário, encaminha seus fragmentos e descarta a cópia remontada. Esse processo adiciona ciclos de CPU e está sujeito a erros.

Alguns aplicativos não dependem da fragmentação, um dos testes mais básicos para verificar o MTU é um ping com uma opção sem fragmento e testar tamanhos de pacotes diferentes: ping endereço ip número de tamanho de df-bit. Se o ping não tiver êxito, corrija a MTU no caminho à medida que a queda ocorrer e causar mais problemas.

Recursos, como Roteamento Baseado em Política e multicaminho de mesmo custo em uma rede com pacotes fragmentados, podem criar problemas de atraso e mais erros, principalmente em altas taxas de dados, induzindo tempos de montagem altos, IDs duplicadas e pacotes corrompidos. Se alguns desses problemas forem identificados, procure resolver essa fragmentação o máximo possível. Um comando para verificar se você tem fragmentos e quaisquer problemas potenciais é show ip traffic:



<#root>

Router#

show ip traffic

IP statistics:

Rcvd: 9875429 total, 14340254 local destination  
0 format errors, 0 checksum errors, 0 bad hop count  
0 unknown protocol, 0 not a gateway  
0 security failures, 0 bad options, 0 with options  
Opts: 0 end, 0 nop, 0 basic security, 0 loose source route  
0 timestamp, 0 extended security, 0 record route  
0 stream ID, 0 strict source route, 0 alert, 0 cipso, 0 ump  
0 other, 0 ignored

**Frag:**

**150 reassembled**

, 0

**timeouts**

,

**0 could not reassemble**

0

**fragmented**

, 600

**fragments**

, 0

**could not fragment**

0 invalid hole

Bcast: 31173 received, 6 sent

Mcast: 0 received, 0 sent

Sent: 15742903 generated, 0 forwarded

Drop: 0 encapsulation failed, 0 unresolved, 0 no adjacency

0 no route, 0 unicast RPF, 0 forced drop, 0 unsupported-addr

0 options denied, 0 source IP address zero

<output omitted>

Na saída acima, as palavras em negrito na seção Frags se referem a:

- Reagrupados: Número de pacotes reagrupados.
- Intervalos: Toda vez que o tempo de remontagem de um fragmento de pacote expira.
- Não foi possível remontar: número de pacotes que não puderam ser remontados.
- Fragmentado: número de pacotes que excedem a MTU e estão sujeitos à fragmentação.
- Fragmentos: Número de blocos nos quais os pacotes foram fragmentados.
- Não foi possível fragmentar: número de pacotes excedendo MTU, mas não foi possível fragmentá-los.

Se a fragmentação for usada e você tiver tempo limite ou não conseguir remontar os contadores, uma forma de corroborar os problemas causados pela plataforma será por quedas de QFP, usando o mesmo comando como explicado mais adiante na seção de quedas: `show platform hardware qfp active statistics drop`. Procure erros como: `TcpBadfrag`, `IpFragErr`, `FragTailDrop`, `ReassDrop`, `ReassFragTooBig`, `ReassTooManyFrag`s, `ReassTimeout` ou outros relacionados. Cada caso pode ter causas diferentes, como não obter todos os fragmentos, duplicados, congestionamento da CPU, entre outros. Novamente, ferramentas úteis para análise posterior e correção potencial podem ser um rastreamento FIA e verificação de configuração.

O TCP oferece o mecanismo MSS (Maximum Segment Size, tamanho máximo de segmento) para resolver esse problema, mas pode induzir latência se o MTU de caminho incorreto, não negociado ou errado for descoberto.

Como o UDP não tem esse mecanismo de fragmentação, você pode confiar na implementação manual do PMTD ou de qualquer solução da camada de aplicação, você pode permitir que eles (quando aplicável) enviem pacotes menores que 576 bytes, que é o MTU efetivo menor para o número de envio de acordo com o RFC1122 em ajudas para evitar a fragmentação.

## Relacionado ao design

Mais do que uma sugestão de solução de problemas, esta seção descreve brevemente mais dois componentes-chave que podem ser adicionados aos problemas de latência e eles exigem uma ampla discussão e análise fora do escopo deste documento.

### Roteamento não ideal

O roteamento não otimizado na rede se refere a uma situação em que os pacotes de dados não estão sendo direcionados pelo caminho mais eficiente ou mais curto disponível em uma rede. Em vez disso, esses pacotes estão tomando uma rota que é menos eficiente, possivelmente resultando em maior latência, congestionamento ou afetando o desempenho da rede. Os IGP's escolhem sempre os melhores caminhos, o que significa menor custo, mas não necessariamente o mais barato ou o caminho de atraso mais baixo (o melhor pode ser aquele com uma largura de banda maior).

O roteamento não ideal pode ocorrer para problemas com protocolos de roteamento; configuração ou qualquer situação, como condições de corrida, alterações dinâmicas (alterações de topologia ou falhas de link), engenharia de tráfego pretendida com base em políticas ou custos da empresa, redundâncias ou failovers (indo para o caminho de backup em determinadas condições), entre outras situações.

Ferramentas como `traceroutes` ou dispositivo de monitoramento podem ajudar a identificar essa situação para fluxos específicos, se esse for o caso, e depende de muitos outros fatores, satisfazer as demandas de aplicativos e a latência mais baixa pode exigir um novo projeto de roteamento ou engenharia de tráfego.

### Quality of Service (QoS)

Ao configurar a qualidade de serviço (QoS), você pode fornecer tratamento preferencial para tipos específicos de tráfego às custas de outros tipos de tráfego. Sem QoS, o dispositivo oferece o serviço de melhor esforço para cada pacote, independentemente do conteúdo ou tamanho do pacote. O dispositivo envia os pacotes sem nenhuma garantia de confiabilidade, limites de atraso ou throughput.

Se a QoS estiver em vigor, torna-se realmente importante identificar se o roteador marca, remarca ou apenas classifica os pacotes, verificar a configuração e mostrar mapa de políticas [nome\_do\_mapa\_de\_políticas | sessão | interface interface\_id] ajuda a entender as classes afetadas por altas taxas, quedas ou pacotes classificados incorretamente.

Implementar a QoS é uma tarefa pesada que requer análise séria e está fora do escopo deste documento, mas é altamente recomendável considerar isso para priorizar aplicativos sensíveis ao tempo e resolver ou evitar muitos problemas de latência e aplicativos.

## Outros problemas de desempenho

Outras condições podem adicionar lentidão, reconexão de sessão ou mau desempenho geral que você precisa verificar, algumas delas são:

### Quedas

Um problema diretamente relacionado ao processamento em um dispositivo é o descarte de pacotes. Você precisa verificar a entrada e a saída do ponto de vista da interface:

```
<#root>
```

```
Router#sh interfaces GigabitEthernet0/0/1
GigabitEthernet0/0/1 is up, line protocol is up
  Hardware is vNIC, address is 0ce0.995d.0000 (bia 0ce0.995d.0000)
  Internet address is 10.10.1.2/24
  MTU 1500 bytes, BW 1000000 Kbit/sec, DLY 10 usec,
    reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  Full Duplex, 1000Mbps, link type is auto, media type is Virtual
  output flow-control is unsupported, input flow-control is unsupported
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:00:19, output 00:08:33, output hang never
  Last clearing of "show interface" counters never

Input queue: 0/375/6788/0 (size/max/drops/flushes); Total output drops: 18263

Queueing strategy: fifo
Output queue: 0/40 (size/max)
  5 minute input rate 114000 bits/sec, 230 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
    193099 packets input, 11978115 bytes, 0 no buffer
    Received 0 broadcasts (0 IP multicasts)
    0 runts, 0 giants, 0 throttles
```

```
1572 input errors
```

```
,  
12 CRC  
, 0 frame,  
1560 overrun  
, 0 ignored  
  0 watchdog, 0 multicast, 0 pause input  
  142 packets output, 11822 bytes, 0 underruns  
  Output 0 broadcasts (0 IP multicasts)  
  0 output errors, 0 collisions, 0 interface resets  
  23 unknown protocol drops  
  0 babbles, 0 late collision, 0 deferred  
  0 lost carrier, 0 no carrier, 0 pause output  
  0 output buffer failures, 0 output buffers swapped out  
Router#
```

No lado de entrada, você tem:

- Quedas de fila de entrada: cada interface possui uma fila de entrada (esse é um buffer de software que pode ser modificado) na qual os pacotes de entrada são colocados para aguardar o processamento pelo Processador de Roteamento (RP). se a taxa de pacotes de entrada colocados na fila de entrada exceder a taxa na qual o RP pode processar os pacotes, você poderá ter descartes incrementais. No entanto, lembre-se de que somente os pacotes de controle e o tráfego "Para nós" são colocados, portanto, se a latência for vista ao passar pelo tráfego, mesmo que você tenha quedas esporádicas, isso não deve ser uma causa.
- Overruns: Isso ocorre quando o hardware do receptor não consegue lidar com os pacotes recebidos para um buffer de hardware porque a taxa de entrada excede a capacidade do receptor de lidar com os dados. Esse número pode indicar um problema com a taxa e o desempenho do roteador, capturar o tráfego somente para essa interface e procurar picos de tráfego. Uma solução comum é ativar o controle de fluxo, mas isso pode adicionar pacotes de retardo. Isso também pode ser uma evidência de gargalos e excesso de assinaturas.
- CRCs: ocorre devido a problemas físicos, verifique o cabeamento, as portas e os SFPs corretamente conectados e o funcionamento adequado.

No lado da saída você tem:

- Quedas de fila de saída: cada interface possui uma fila de saída na qual são colocados os pacotes de saída a serem enviados na interface. Às vezes, a taxa de pacotes de saída colocados na fila de saída pelo RP excede a taxa na qual a interface pode enviar os pacotes. Isso pode causar problemas de desempenho e problemas de latência se não houver QoS em vigor. Caso contrário, você pode ter esse número aumentando por causa de uma determinada política aplicada e aconselhar a verificação ou implementação da configuração de QoS para proteger e garantir o tráfego intencional ou crítico.

Por fim, descartes no QFP estão diretamente relacionados ao alto processamento que pode

causar latência, verifique via show platform hardware qfp active statistics drop:

```
<#root>
```

```
Router#
```

```
show platform hardware qfp active statistics drop
```

```
Last clearing of QFP drops statistics : never
```

Global Drop Stats	Packets	Octets
Disabled	2	646
Ipv4NoAdj	108171	6706602
Ipv6NoRoute	10	560

As causas dependem do código, o rastreamento FIA ajuda a corroborar ou descartar se o tráfego afetado pela latência for descartado nesse ponto.

## Retransmissão TCP

A retransmissão de TCP é um sintoma ou pode ser uma consequência devido a um problema subjacente, como a perda de pacotes. Esse problema pode causar lentidão e mau desempenho no aplicativo.

O Protocolo de Controle de Transmissão (TCP - Transmission Control Protocol) usa um temporizador de retransmissão para garantir a entrega de dados na ausência de qualquer feedback do receptor de dados remoto. A duração desse temporizador é conhecida como RTO (retransmission timeout). Quando o temporizador de retransmissão expira, o emissor retransmite o segmento mais antigo que não foi confirmado pelo receptor TCP e o RTO aumenta.

Algumas retransmissões não podem ser completamente eliminadas, se forem mínimas, não podem refletir um problema. No entanto, como você pode inferir, mais retransmissão vista, mais latência na sessão TCP e precisa ser endereçada.

A captura de pacotes analisada no Wireshark pode corroborar o problema como o próximo exemplo:

The image shows a Wireshark packet capture of a TCP session. The packet list pane displays several segments, with some highlighted in blue to indicate retransmissions. The packet details pane for a selected packet shows the following information:

- Protocol: TCP
- Sequence: 1000000
- Length: 1000
- Flags: R (Retransmission)
- Window: 65535
- Checksum: 0x1234
- Source: 10.0.0.1
- Destination: 10.0.0.2

The packet details pane also shows a note: "Retransmission" and "This frame is a (suspected) retransmission".

Se houver retransmissões, use o mesmo método de captura na direção de entrada e saída do roteador para verificar todos os pacotes enviados e recebidos. Naturalmente, fazer isso em cada salto pode representar um enorme esforço, de modo que a análise detalhada da captura é necessária para o TCP, observando TTLs, vezes de quadros anteriores no mesmo fluxo de TCP para entender de que direção (servidor ou cliente) você tem esse atraso ou falta de resposta para direcionar sua solução de problemas.

## Excesso de demanda e gargalos

O excesso de assinaturas ocorre quando os recursos necessários (largura de banda) são maiores que os disponíveis reais. Os comandos para identificar se você tem esse problema em um roteador já foram abordados na seção anterior.

Como consequência dessa situação, os gargalos podem ocorrer quando os fluxos de tráfego ficam mais lentos devido à largura de banda ou capacidade de hardware insuficiente. É importante identificar se isso acontece em pouco tempo ou se é uma situação de longo prazo para aplicar soluções.

Não há nenhum conselho específico para resolvê-lo, mas algumas das opções são equilibrar o tráfego para diferentes plataformas, segmentar a rede ou atualizar para dispositivos mais robustos com base nas necessidades atuais e na análise de crescimento futuro.

## Informações Relacionadas

- [Operações de eco ICMP de SLAs IP](#)
- [Solução de problemas de memória](#)
- [Solucionar problemas com o recurso de rastreamento de pacote de caminho de dados do Cisco IOS-XE](#)
- [Solucione problemas de quedas de pacotes nos roteadores de serviço ASR 1000 Series.](#)
- [Informações relacionadas à QoS](#)
- [Configuração de QoS em roteadores](#)
- [Suporte técnico e downloads da Cisco](#)

## Sobre esta tradução

A Cisco traduziu este documento com a ajuda de tecnologias de tradução automática e humana para oferecer conteúdo de suporte aos seus usuários no seu próprio idioma, independentemente da localização.

Observe que mesmo a melhor tradução automática não será tão precisa quanto as realizadas por um tradutor profissional.

A Cisco Systems, Inc. não se responsabiliza pela precisão destas traduções e recomenda que o documento original em inglês ([link fornecido](#)) seja sempre consultado.