Solucione problemas de cenários com fixação Null0 e MSS

Contents

Introdução

Este documento descreve as implicações de se ter o ajuste de Tamanho Máximo de Segmento (MSS - Maximum Segment Size) e rotas estáticas que apontam para Nulo 0 no Catalyst 9K.

Pré-requisitos

Requisitos

A Cisco recomenda que você conheça estes tópicos:

- Conhecimento conceitual sobre ajuste de TCP e MSS
- Compreensão da plataforma do Cisco Catalyst 9K para encaminhamento e depurações do plano de controle.

Plataformas suportadas

Este documento se aplica a todas as plataformas Catalyst 9K que executam o Cisco IOS® XE 17.3.x e posterior.

Componente usado

As informações neste documento são baseadas nestas versões de software e hardware:

- Catalyst 9300 Series Switches executando a versão IOS-XE 17.3.4
- Catalyst 9400 Series Switches executando a versão IOS-XE 17.3.4
- IXIA para geração de tráfego

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.

Abordagem de solução de problemas

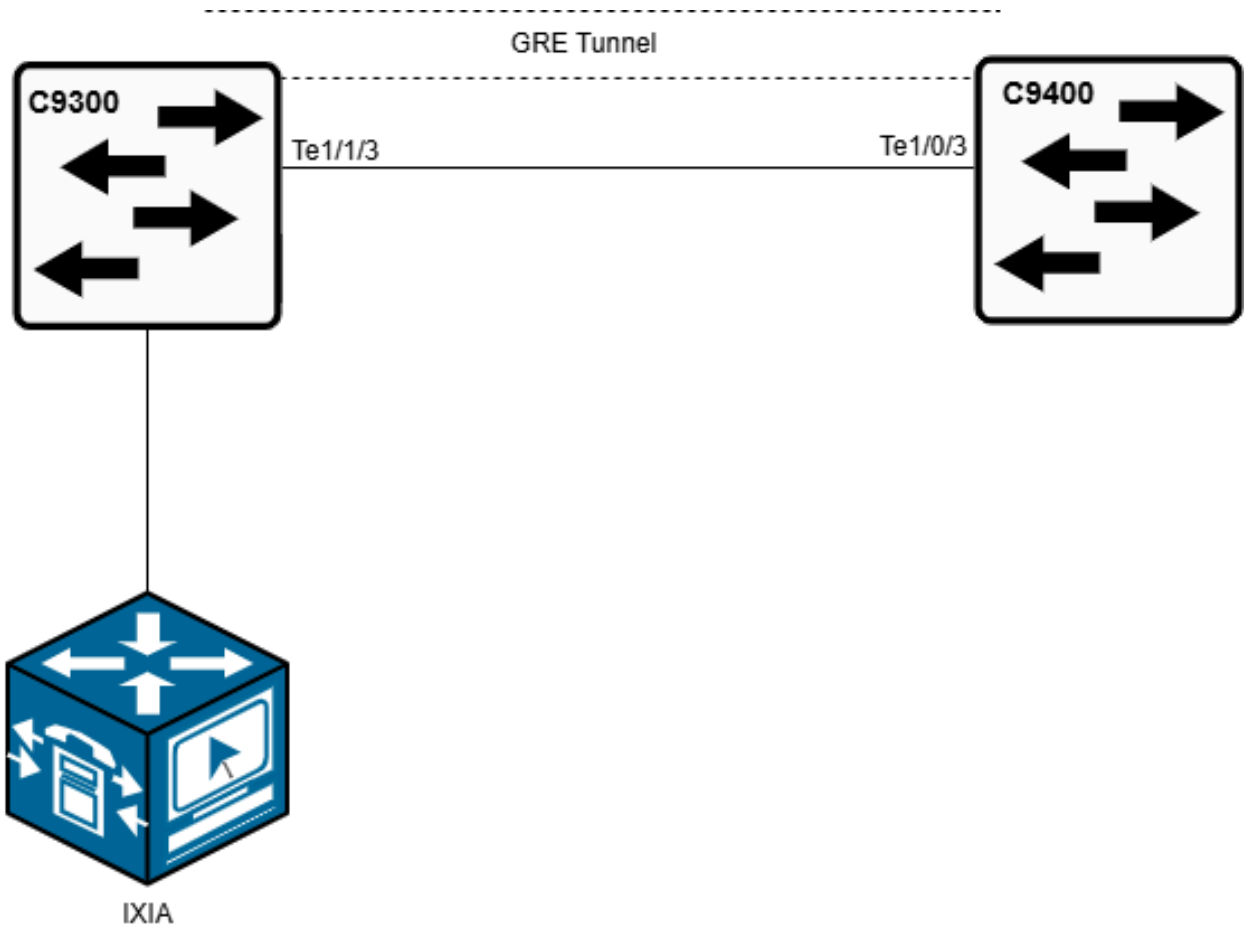
Topologia

A configuração consiste em switches C9000 com um gerador de tráfego para reproduzir o problema. Testes incluídos para isolamento adicional:

Condição 1: Sem 'Null0' ou 'MSS adjust'
Condição 2: Com uma rota estática apontando para Null0, nenhum ajuste de MSS
Condição 3: Ajuste Null0 e MSS habilitado

Versões de software e hardware

- Catalyst 9300 e 9400 executando o Cisco IOS XE versão 17.3.4
- IXIA para geração de tráfego



Requisitos de configuração

- Nenhum 'ip tcp adjust-mss' e nenhuma 'rota null0' configurados
- Com apenas 'null0 route' configurado
- Com 'ip tcp adjust-mss' e 'null0 route' configurados
  'ip tcp adjust-mss value' (valor menor que a Unidade máxima de transmissão (MTU)) [Na interface de túnel ou na interface virtual do switch (SVI) (Entrada)]
  'ip route X.X.X.X X.X.X.X Null0' (Rotas estáticas apontando para Null0)

Com base nas condições descritas, você observa a conectividade intermitente com peers BGP (Border Gateway Protocol) diretamente conectados e com SVIs configurados no mesmo dispositivo ou em peers diretamente conectados. Há também um aumento consistente nos contadores de queda na fila de Encaminhamento do software (SW) durante a execução de comandos e depurações de Política de Plano de Controle (CoPP). A investigação mostra que o tráfego destinado a Null0 é direcionado para a CPU. Esse comportamento interrompeu o protocolo BGP, impedindo a conclusão do handshake triplo do TCP. Além disso, os pings para os endereços IP do SVI configurados no switch falharam.

Cenários

Caso 1. Sem 'Null0' ou 'MSS Adjust'

Se nem 'ip tcp adjust-mss' nem uma 'rota nula' estiverem configurados, o contador de descarte na fila de encaminhamento de SW permanecerá em '0' após o tráfego gerado do IXIA, como esperado.

Consulte estes registros:

```
Cat-9400-1# Show platform hardware fed active qos queue stats internal cpu policer
CPU Queue Statistics
============================================================================================
(default) (set) Queue Queue
QId PlcIdx Queue Name Enabled Rate Rate Drop(Bytes) Drop(Frames)
--------------------------------------------------------------------------------------------
14 13 Sw forwarding Yes 1000 200 0 0>>>>>>>>>>>>>>>>>>>>>>>>>>>>No increment
```

Caso 2. Com uma rota estática apontando para Null0, Sem ajuste de MSS

Com apenas a rota Null0 configurada, o contador de queda na fila de encaminhamento de SW permanece em '0' após o tráfego gerado do IXIA, como esperado.

Consulte estes registros:

```
Cat-9400-1# Show platform hardware fed active qos queue stats internal cpu policer
CPU Queue Statistics
============================================================================================
(default) (set) Queue Queue
QId PlcIdx Queue Name Enabled Rate Rate Drop(Bytes) Drop(Frames)
--------------------------------------------------------------------------------------------
14 13 Sw forwarding Yes 1000 200 0 0>>>>>>>>>>>>>>>>>>>>>>>>>>>>No increment
```

Caso 3. 'Null0' e 'MSS Adjust' Ativados

```
With both "ip tcp adjust-mss" and a "null route" configured:
```

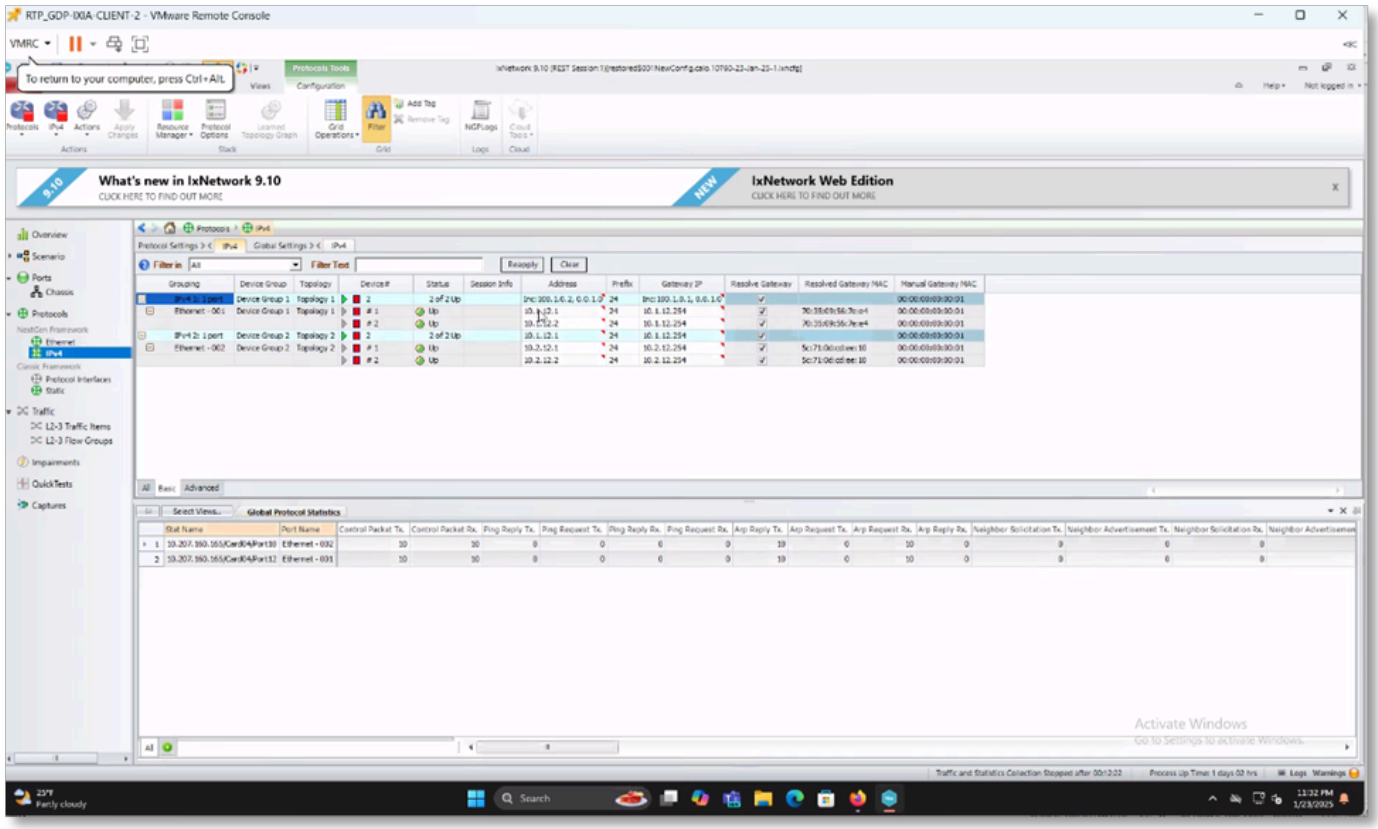
```
Configuration:

On Cat 9300:
Cat-9300-1#show run interface twoGigabitEthernet 1/0/1
interface TwoGigabitEthernet1/0/1 (Interface connected to IXIA)
no switchport
ip address 10.1.12.xx 255.255.255.0
end
Cat-9300-1#show run interface tenGigabitEthernet 1/1/3
interface TenGigabitEthernet1/1/3 (Physical interface connected to C9400)
```

```
no switchport
mtu 9000
ip address 203.63.147.xx 255.255.255.0
no ip redirects
no ip unreachables
ip mtu 1500
load-interval 30
end

Cat-9300-1#show run interface tunnel421
interface Tunnel421
description Tunnel 421 to Scrubbing Center - SYD EDGE 1 and 2 - AR1 Tunnel 30
ip address 10.88.178.xx 255.255.255.0
ip mtu 1470
load-interval 30
Cisco Confidential
keepalive 10 3
tunnel source 203.63.147.xx
tunnel destination 203.63.147.xx
end

On cat 9400:
Cat-9400-1#show run interface tenGigabitEthernet 1/0/3
interface TenGigabitEthernet1/0/3 (Interface connected to C9300)
description CN,ISP,S1 AAPT, Superloop Circuit ID SID565199 - AAPT Circuit ID 5804194
no switchport
mtu 9000
ip address 203.63.147.xx 255.255.255.0
no ip redirects
no ip unreachables
ip mtu 1500
load-interval 30
end

interface Tunnel421
description Tunnel 421 to Scrubbing Center - SYD EDGE 1 and 2 - AR1 Tunnel 30
ip address 10.88.178.xx 255.255.255.0
ip mtu 1470
ip tcp adjust-mss 500>>>>>>>>>>>>
load-interval 30
keepalive 10 3
tunnel source 203.63.147.xx
tunnel destination 203.63.147.xx
end

Null0 Routes:
ip route 10.2.12.xx 255.255.255.255 null0>>>>>>>>Destination IP is of IXIA connected to 9300

Cat-9400-1#show ip route
Gateway of last resort is 203.63.147.xx to network 0.0.0.0
S* 0.0.0.0/0 [1/0] via 203.63.147.xx
10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
S 10.2.12.0/24 [1/0] via 192.168.12.xx
S 10.2.12.xx/32 is directly connected, Null0
C 10.88.178.0/24 is directly connected, Tunnel421
L 10.88.178.xx/32 is directly connected, Tunnel421
```

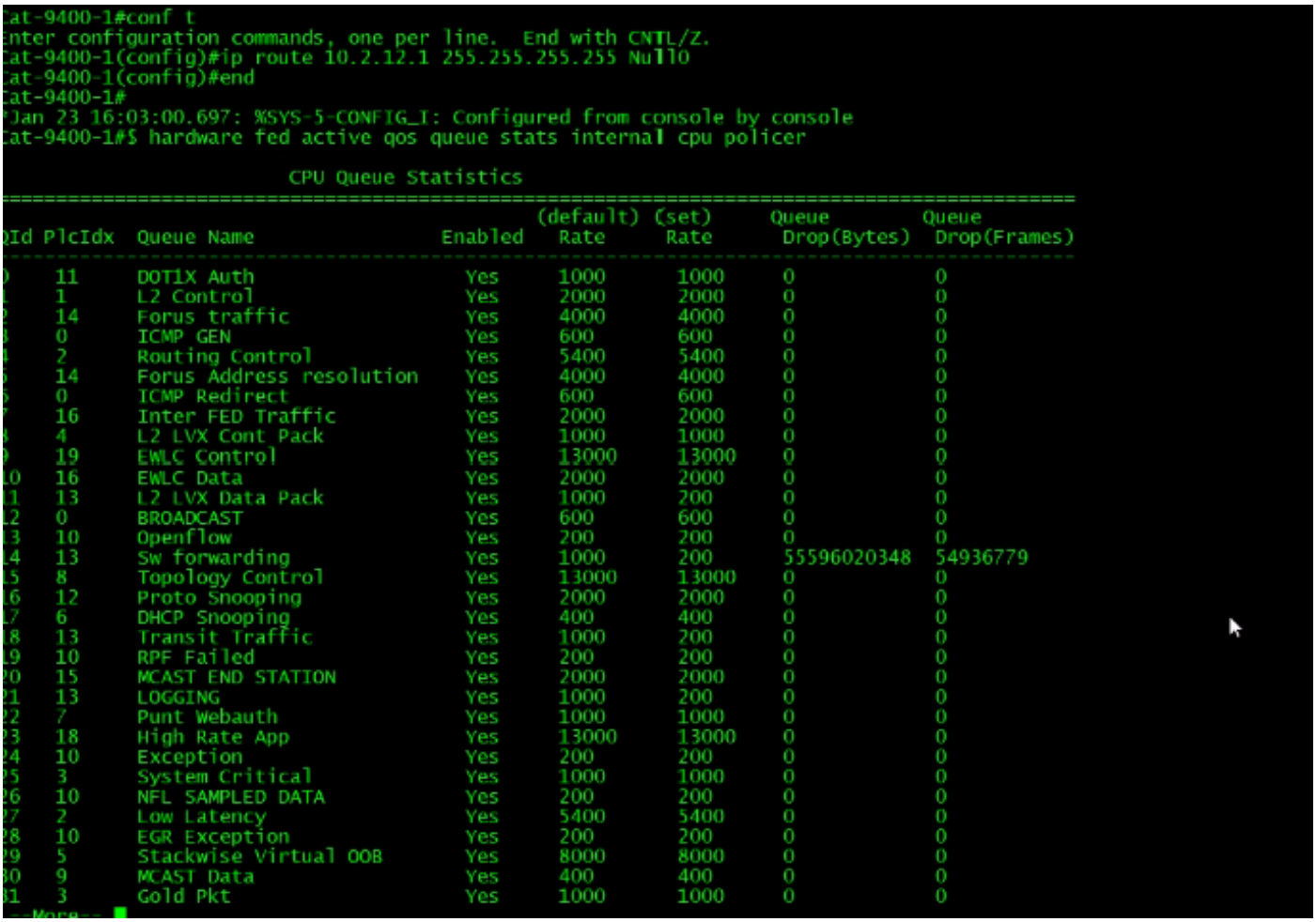
Depois que as rotas Null0 e o MSS ajustam a configuração na interface de túnel de entrada do C9400, o tráfego foi gerado do IXIA e o contador de queda incrementa para a Identidade da fila

da CPU (QID) 14, como mostrado na imagem a seguir.

IXIA



Saída CoPP do C9400:



```
Cat-9400-1#conf t
Enter configuration commands, one per line.  End with CNTL/Z.
Cat-9400-1(config)#ip route 10.2.12.1 255.255.255.255 Null0
Cat-9400-1(config)#end
Cat-9400-1#
Jan 23 16:03:00.697: %SYS-5-CONFIG_I: Configured from console by console
Cat-9400-1#$ hardware fed active qos queue stats internal cpu policer

                    CPU Queue Statistics
============================================================================================
                                                (default) (set)     Queue         Queue
QId PlcIdx  Queue Name                Enabled   Rate      Rate      Drop(Bytes)   Drop(Frames)
--------------------------------------------------------------------------------------------
0    11     DOT1X Auth                Yes      1000      1000      0             0
1    1      L2 Control                Yes      2000      2000      0             0
2    14     Forus traffic             Yes      4000      4000      0             0
3    0      ICMP GEN                  Yes      600       600       0             0
4    2      Routing Control           Yes      5400      5400      0             0
5    14     Forus Address resolution  Yes      4000      4000      0             0
6    0      ICMP Redirect             Yes      600       600       0             0
7    16     Inter FED Traffic         Yes      2000      2000      0             0
8    4      L2 LVX Cont Pack          Yes      1000      1000      0             0
9    19     EWLC Control              Yes      13000     13000     0             0
10   16     EWLC Data                 Yes      2000      2000      0             0
11   13     L2 LVX Data Pack          Yes      1000      200       0             0
12   0      BROADCAST                 Yes      600       600       0             0
13   10     Openflow                  Yes      200       200       0             0
14   13     Sw forwarding             Yes      1000      200       55596020348   54936779
15   8      Topology Control          Yes      13000     13000     0             0
16   12     Proto Snooping            Yes      2000      2000      0             0
17   6      DHCP Snooping             Yes      400       400       0             0
18   13     Transit Traffic           Yes      1000      200       0             0
19   10     RPF Failed                Yes      200       200       0             0
20   15     MCAST END STATION         Yes      2000      2000      0             0
21   13     LOGGING                   Yes      1000      200       0             0
22   7      Punt Webauth              Yes      1000      1000      0             0
23   18     High Rate App             Yes      13000     13000     0             0
24   10     Exception                 Yes      200       200       0             0
25   3      System Critical           Yes      1000      1000      0             0
26   10     NFL SAMPLED DATA          Yes      200       200       0             0
27   2      Low Latency               Yes      5400      5400      0             0
28   10     EGR Exception             Yes      200       200       0             0
29   5      Stackwise Virtual OOB     Yes      8000      8000      0             0
30   9      MCAST Data                Yes      400       400       0             0
31   3      Gold Pkt                  Yes      1000      1000      0             0
--More--
```

```
Cat-9400-1# show platform hardware fed active qos queue stats internal cpu policer
============================================================================================
(default) (set) Queue Queue
QId PlcIdx Queue Name Enabled Rate Rate Drop(Bytes) Drop(Frames)
--------------------------------------------------------------------------------------------
14 13 Sw forwarding Yes 1000 200 3252568000 3214000>>>>>> Drops increasing in this Queue

Cat-9400-1# show platform hardware fed active qos queue stats internal cpu policer

                        CPU Queue Statistics

============================================================================================

                                                (default) (set)     Queue         Queue

QId PlcIdx  Queue Name                Enabled   Rate      Rate      Drop(Bytes)   Drop(Frames)

--------------------------------------------------------------------------------------------

0    11     DOT1X Auth                Yes      1000      1000      0             0

1    1      L2 Control                Yes      2000      2000      0             0

2    14     Forus traffic             Yes      4000      4000      0             0

3    0      ICMP GEN                  Yes      600       600       0             0

4    2      Routing Control           Yes      5400      5400      0             0
```

```
5    14     Forus Address resolution    Yes     4000      4000     0            0

6    0      ICMP Redirect               Yes     600       600      0            0

7    16     Inter FED Traffic           Yes     2000      2000     0            0

8    4      L2 LVX Cont Pack            Yes     1000      1000     0            0

9    19     EWLC Control                Yes     13000     13000    0            0

10   16     EWLC Data                   Yes     2000      2000     0            0

11   13     L2 LVX Data Pack            Yes     1000      200      0            0

12   0      BROADCAST                   Yes     600       600      0            0

13   10     Openflow                    Yes     200       200      0            0

14   13     Sw forwarding               Yes     1000      200      40147794808 39671734>>>>>>With MSS a

15   8      Topology Control            Yes     13000     13000    0            0

16   12     Proto Snooping              Yes     2000      2000     0            0

17   6      DHCP Snooping               Yes     400       400      0            0
```

Explicação de rotas estáticas Null0 e fixação de MSS

De acordo com a teoria, para lidar com o tráfego indesejado, como o tráfego de broadcast, ou bloquear o acesso a sub-redes específicas, uma opção é configurar uma rota estática que direcione o tráfego para Null0. Isso faz com que o roteador descarte qualquer tráfego destinado a essa rede.

Comando para Null0

```
ip route <destination-network> <subnet-mask> null 0

For an example:

ip route 10.2.12.xx 255.255.255.255 null0>>>>>>Destination IP is of IXIA connected to 9300
```

A sintaxe null 0 garante que 10.2.12.1/32 não seja encaminhado para nenhum lugar. O que significa que todo o tráfego destinado à rede destino é descartado (descartado) em Null0.

TCP MSS

Por outro lado, o TCP MSS Adjustment:

O ajuste de MSS modifica o MSS para pacotes TCP. Quando ocorre uma incompatibilidade de MTU, geralmente entre dispositivos com configurações de MTU diferentes ou através de túneis como VPNs, os pacotes podem ser fragmentados.

A fragmentação não é desejável para o tráfego TCP porque pode levar à perda de pacotes ou à degradação do desempenho. O aperto de MSS resolve esse problema ajustando o tamanho dos segmentos TCP, garantindo que os pacotes sejam pequenos o suficiente para caber dentro do MTU do caminho e, portanto, evita a fragmentação. Quando o ajuste de MSS é aplicado a interfaces de túnel e SVIs com um valor definido como 1360 para conexões TCP, ele garante que o tamanho do segmento seja menor que o MTU do caminho, o que evita a fragmentação.

Cenário ideal

Null0 é uma interface virtual de "buraco negro" que descarta qualquer tráfego direcionado a ela. É útil para evitar loops de roteamento ou tráfego indesejado.
O ajuste de TCP MSS é um comando que garante que os segmentos TCP sejam pequenos o suficiente para evitar a fragmentação ao passar por dispositivos ou túneis com MTUs menores.

Condição

Embora esses dois recursos sejam geralmente usados para finalidades diferentes, ambos podem desempenhar um papel em um projeto geral de rede para gerenciar o fluxo de tráfego, evitar a fragmentação e otimizar o desempenho. No entanto, nos switches Catalyst 9K, o uso conjunto do ajuste Null0 e MSS pode gerar conflitos, sobrecarregar a CPU e sobrecarregar a política de CoPP.

Verificação

```
Show platform hardware fed active qos queue stats internal cpu policer
Identify the QID where the drop counters increments. After finding the QID (for example, QID 14), run t
#debug platform software fed switch active punt packet-capture set-filter "fed.queue == 14"
#debug platform software fed switch active punt packet-capture start
#debug platform software fed switch active punt packet-capture stop
#show platform software fed switch active punt packet-capture brief
#show platform software fed switch active punt packet-capture detailed
```

Usando os comandos debug, verifique os logs no próximo formato para identificar o endereço IP dos punts dos invasores na CPU, mesmo com as rotas Null0 configuradas:

```
------ Punt Packet Number: XX, Timestamp: 2024/12/14 12:54:57.508 ------
interface : physical: [if-id: 0x00000000], pal: Tunnel411 [if-id: 0x000000d2]
metadata : cause: 11 [For-us data], sub-cause: 1, q-no: 14, linktype: MCP_LINK_TYPE_IP [1]
ether hdr : Partial ether header, ethertype: 0x0800 (IPv4)
Cisco Confidential
ipv4 hdr : dest ip: XX.XX.XX.XX, src ip: XX.XX.XX.XX
```

```
ipv4 hdr : packet len: 44, ttl: 242, protocol: 6 (TCP)
tcp hdr : dest port: 777, src port: 41724
```

Debugs

```
Cat-9400-1# debug platform software fed active punt packet-capture set-filter "fed.queue == 14"
Filter setup successful. Captured packets will be cleared

Cat-9400-1#debug platform software fed active punt packet-capture start
Punt packet capturing started.

Cat-9400-1#debug platform software fed active punt packet-capture stop
Punt packet capturing stopped. Captured 4096 packet(s)

Cat-9400-1#show platform software fed active punt packet-capture brief
Total captured so far: 4096 packets. Capture capacity : 4096 packets
Capture filter : "fed.queue == 14"
------ Punt Packet Number: 1, Timestamp: 2025/01/23 16:16:54.978 ------
interface : physical: [if-id: 0x00000000], pal: Tunnel421 [if-id: 0x0000002e]
metadata : cause: 11 [For-us data], sub-cause: 1, q-no: 14, linktype: MCP_LINK_TYPE_IP [1]
ether hdr : Partial ether header, ethertype: 0x0800 (IPv4)
ipv4 hdr : dest ip: 10.2.12.xx, src ip: 10.1.12.xx >>>10.2.12.xx is IXIA
ipv4 hdr : packet len: 1006, ttl: 63, protocol: 6 (TCP)
tcp hdr : dest port: 60, src port: 60
------ Punt Packet Number: 2, Timestamp: 2025/01/23 16:16:54.978 ------
interface : physical: [if-id: 0x00000000], pal: Tunnel421 [if-id: 0x0000002e]
metadata : cause: 11 [For-us data], sub-cause: 1, q-no: 14, linktype: MCP_LINK_TYPE_IP [1]
ether hdr : Partial ether header, ethertype: 0x0800 (IPv4)
ipv4 hdr : dest ip: 10.2.12.xx, src ip: 10.1.12.xx >>>10.2.12.xx is IXIA
ipv4 hdr : packet len: 1006, ttl: 63, protocol: 6 (TCP)
tcp hdr : dest port: 60, src port: 60
------ Punt Packet Number: 3, Timestamp: 2025/01/23 16:16:54.978 ------
interface : physical: [if-id: 0x00000000], pal: Tunnel421 [if-id: 0x0000002e]
metadata : cause: 11 [For-us data], sub-cause: 1, q-no: 14, linktype: MCP_LINK_TYPE_IP [1]
ether hdr : Partial ether header, ethertype: 0x0800 (IPv4)
ipv4 hdr : dest ip: 10.2.12.xx, src ip: 10.1.12.xx >>>10.2.12.xx is IXIA
Cisco Confidential
ipv4 hdr : packet len: 1006, ttl: 63, protocol: 6 (TCP)
tcp hdr : dest port: 60, src port: 60
```

Conclusão

Para evitar que as filas da CPU sejam sobrecarregadas pelo tráfego indesejado e afetem a comunicação TCP/Secure Shell (SSH), bloqueie esses endereços IP antes que eles alcancem os switches Catalyst 9K ou remova o ajuste de MSS na entrada.

Normalmente, o pacote SYN sincroniza TCP para a fila da CPU. O MSS é uma opção no cabeçalho TCP que indica o tamanho máximo de segmento que o receptor pode aceitar, exceto cabeçalhos TCP/IP. Ele é geralmente definido para o handshake triplo, especificamente no pacote SYN.

Para resolver esse problema, bloqueie geograficamente os IPs mal-intencionados no gateway de RADWARE/segurança para evitar que a fila do vigilante da CPU fique sobrecarregada e estabilize o peering de BGP e as conexões TCP.

Resolução

Depois que os IPs mal-intencionados foram bloqueados no gateway de segurança/Radware com êxito, o tráfego parou de sobrecarregar a fila da CPU.

Informações Relacionadas

- https://www.cisco.com/c/en/us/support/docs/ip/transmission-control-protocol-tcp/222338-troubleshoot-tcp-slowness-issues-due-to.html
- 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.