Troubleshoot Scenarios with Null0 and MSS Clamping

Contents

Introduction

Prerequisites

- Requirements
- Supported Platforms
- Component Used

Troubleshooting Approach

- Topology
- Software and Hardware Versions

Configuration Requirements

Scenarios

- Case 1. Without 'Null0' or 'MSS Adjust'
- Case 2. With a Static Route Points to Null0, No MSS Adjust
- Case 3. Both 'Null0' and 'MSS Adjust' Enabled
- IXIA

Explanation of Null0 Static Routes and MSS Clamping

- Command for Null0
- TCP MSS
- Ideal Scenario

Condition

Verification

Debugs

Conclusion

Resolution

Related Information

Introduction

This document describes the implications of having Maximum Segment Size (MSS) adjustment and static routes pointing to Null 0 on Catalyst 9K.

Prerequisites

Requirements

Cisco recommends you have knowledge of these topics:

- Conceptual knowledge on TCP and MSS adjust
- Platform understanding of Cisco Catalyst 9K for control plane forwarding and debugs.

Supported Platforms

This document is applicable for all Catalyst 9K platform running Cisco IOS® XE 17.3.x and later.

Component Used

The information in this document is based on these software and hardware versions:

- Catalyst 9300 series switches running IOS-XE 17.3.4 version
- Catalyst 9400 series switches running IOS-XE 17.3.4 version
- IXIA for generating traffic

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.

Troubleshooting Approach

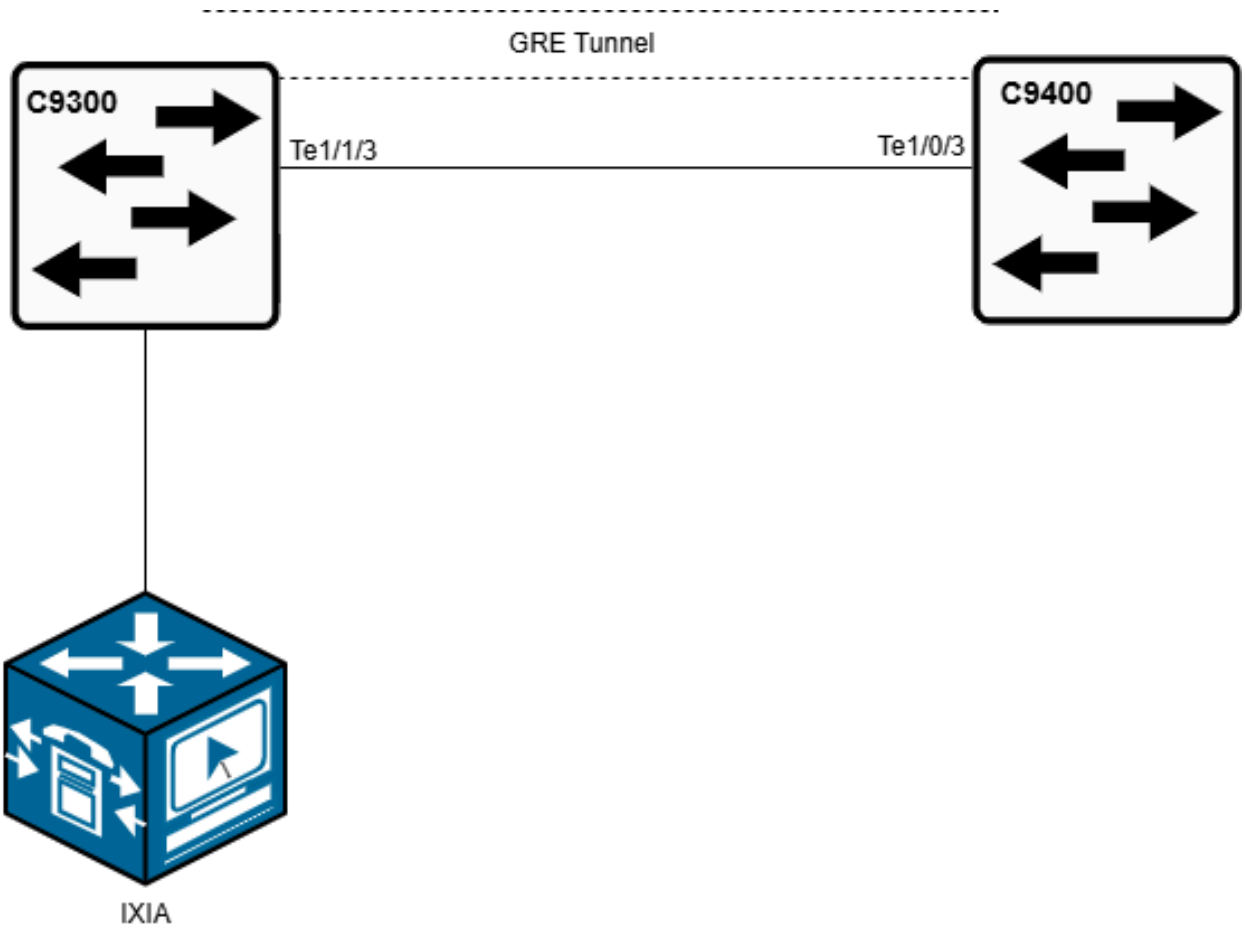
Topology

The setup consists of C9000 switches with a traffic generator in order to reproduce the issue. Tests included for further isolation:

Condition 1: Without 'Null0' or 'MSS adjust'
Condition 2: With a static route pointing to Null0, no MSS adjust
Condition 3: Both Null0 and MSS adjust enabled

Software and Hardware Versions

- Catalyst 9300 and 9400 running Cisco IOS XE 17.3.4 version
- IXIA for generating traffic



Configuration Requirements

- No 'ip tcp adjust-mss' and no 'null0 route' configured
- With only 'null0 route' configured
- With 'ip tcp adjust-mss' and 'null0 route' configured
  'ip tcp adjust-mss value' (value less than Maximum Transmission Unit (MTU)) (On Tunnel interface or Switch Virtual Interface (SVI) (Ingress))
  'ip route X.X.X.X X.X.X.X Null0' (Static routes pointing to Null0)

Based on the conditions described, you observe intermittent connectivity to directly connected Border Gateway Protocol (BGP) peers and to SVIs configured on the same device or on directly connected peers. There is also a consistent increase in drop counters in the software (SW) Forwarding queue while running Control Plane Policing (CoPP) commands and debugs. Investigation shows that traffic intended for Null0 is instead directed to the CPU. This behavior disrupted the BGP protocol by preventing the TCP 3-way handshake completion. Additionally, pings to the SVI IP addresses configured on the switch failed.

Scenarios

Case 1. Without 'Null0' or 'MSS Adjust'

If neither 'ip tcp adjust-mss' nor a 'null route' is configured, the drop counter in the SW forwarding queue remains at '0' after traffic generated from IXIA, as expected.

Refer these logs:

```
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
```

Case 2. With a Static Route Points to Null0, No MSS Adjust

With only the Null0 route configured, the drop counter in the SW forwarding queue remains at '0' after traffic generated from IXIA, as expected.

Refer these logs:

```
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
```

Case 3. Both 'Null0' and 'MSS Adjust' Enabled

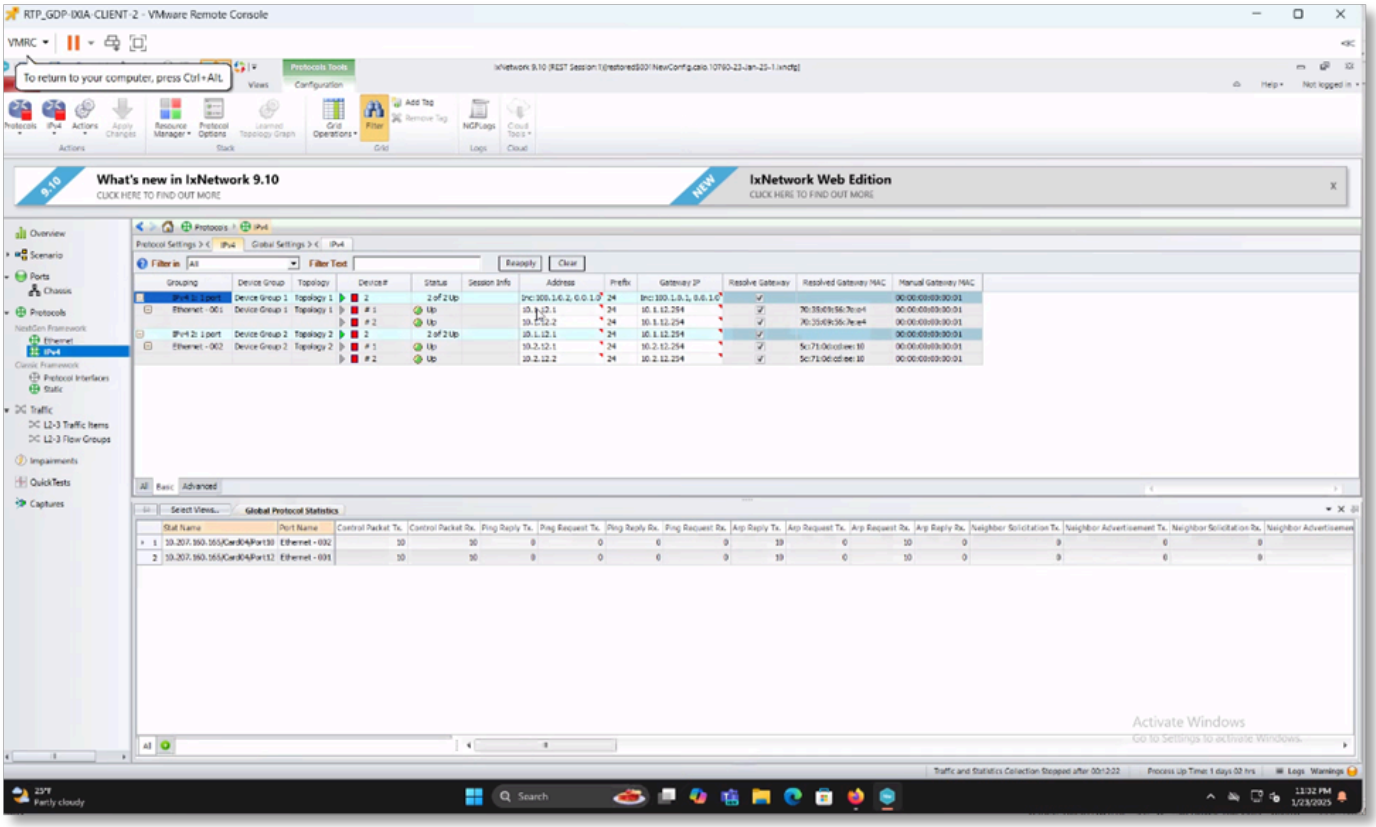
```
With both "ip tcp adjust-mss" and a "null route" configured:
```

```
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
```

After Null0 routes and MSS adjust configuration on the ingress tunnel interface of the C9400, traffic was generated from IXIA, and the drop counter increments for CPU queue Identity (QID) 14, as shown in the next image.

IXIA



C9400 CoPP output:

```
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
```

Explanation of Null0 Static Routes and MSS Clamping

According to theory, in order to handle unwanted traffic such as broadcast traffic or block access to specific subnets, one option is to set up a static route that directs traffic to Null0. This causes the router to discard

any traffic intended for that network.

Command for 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
```

Null 0 syntax ensures that the 10.2.12.1/32 not be forwarded anywhere. Which means that any traffic destined for destination network is discarded (dropped) at Null0.

TCP MSS

On the other hand, TCP MSS Adjustment:

MSS adjustment modifies the MSS for TCP packets. When an MTU mismatch occurs—often between devices with different MTU settings or through tunnels like VPNs—packets can be fragmented.

Fragmentation is undesirable for TCP traffic because it can lead to packet loss or performance degradation. MSS clamping addresses this issue by adjusting the size of TCP segments, ensuring that packets are small enough to fit within the path MTU, and thus prevents fragmentation. When MSS adjustment is applied to tunnel interfaces and SVIs with a value set to 1360 for TCP connections, it ensures that the segment size is smaller than the path MTU, which prevents fragmentation.

Ideal Scenario

Null0 is a virtual 'black hole' interface that drops any traffic directed toward it. It is useful to prevent routing loops or unwanted traffic.
TCP MSS adjust is a command that ensures TCP segments are small enough to avoid fragmentation when passing through devices or tunnels with smaller MTUs.

Condition

While these two features are generally used for different purposes, they can both play a role in an overall network design in order to manage traffic flow, avoid fragmentation, and optimize performance. However, on Catalyst 9K switches, using both Null0 and MSS adjustment together can lead to conflicts, overload the CPU and overwhelms the CoPP policy.

Verification

```
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
```

Using the debug commands, check the logs in the next format in order to identify the IP address of the attackers punts on the CPU, even with the Null0 routes configured:

```
------ 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
```

Conclusion

In order to prevent CPU queues from being overwhelmed by unwanted traffic and affecting TCP/Secure Shell (SSH) communication, block these IP addresses before they reach the Catalyst 9K switches or remove MSS adjustment on ingress.

Typically, the TCP synchronize (SYN) packet punts to the CPU queue. MSS is an option in the TCP header that indicates the maximum segment size the receiver can accept, except TCP/IP headers. It is usually set for the 3-way handshake, specifically in the SYN packet.

In order to resolve this issue, geo-block the malicious IPs on the RADWARE/Security Gateway to prevent the CPU policer queue from becoming overwhelmed and stabilize BGP peering and TCP connections.

Resolution

Once malicious IPs blocked on the Radware/security gateway successfully, traffic stopped overwhelming the CPU queue.

Related Information

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