Verify Control Plane Policing Violations on Nexus Platforms

Contents

Introduction

This document describes details about Control Plane Policing (CoPP) on Cisco Nexus switches and its relevant impact on non-default class violations.

Prerequisites

Cisco recommends that you understand basic information with regard to Control Plane Policing (CoPP), its guidelines and limitations, and general configuration, as well as Quality-of-Service (QoS) policing (CIR) functionality. For more information about this feature, refer to the applicable documents:

- Cisco Nexus 9000 Series NX-OS Security Configuration Guide, Release 10.2(x)
- CoPP on Nexus 7000 Series Switches
- Cisco Nexus 9000 Series NX-OS Quality of Service Configuration Guide, Release 10.2(x)

Requirements

There are no specific requirements for this document.

Components Used

This document is not restricted to specific software and hardware requirements.

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.

Background Information

Control plane traffic is redirected toward the supervisor module by redirection access control lists (ACLs) programmed to punt the matched traffic that passes through two layers of protection, the hardware rate-limiters and CoPP. Any disruptions or attacks to the supervisor module, if left unchecked, can result in serious network outages; thus CoPP is there to serve as a protection mechanism. If there is instability at the control plane level, it is important to check CoPP, because abnormal traffic patterns created from loops or floods, or rogue devices can tax and prevent the supervisor from processing legitimate traffic. Such attacks, which can be perpetrated either inadvertently by rogue devices or maliciously by attackers, typically involve high rates of traffic destined to the supervisor module or the CPU.

Control Plan Policing (CoPP) is a feature that classifies and polices all packets received over the in-band (front panel) ports destined to the router address or that require any supervisor involvement. This feature allows a policy map to be applied to the control plane. This policy map looks like a normal quality of service (QoS) policy and is applied to all traffic that enters the switch from a non-management port. Protection of the supervisor module by policing allows the switch to mitigate floods of traffic that go beyond the committed input rate (CIR) for each class by the discard of packets to prevent the switch from being overwhelmed and thus an impact on performance.

It is important to monitor CoPP counters continuously and to justify them, which is the purpose of this document. CoPP violations, if left unchecked, can prevent the control plane from the process of genuine traffic on the associated affected class. CoPP configuration is a fluid and on-going process that must respond to the network and infrastructure requirements. There are three default system policies for CoPP. By default, Cisco recommends the use of the default policy `strict` as the initial start point and is used as the basis for this document.

CoPP only applies to in-band traffic received through the front panel ports. The out-of-band management port (mgmt0) is not subject to CoPP. The Cisco NX-OS device hardware performs CoPP on a per-forwarding-engine basis. Therefore, choose rates so the aggregate traffic does not overwhelm the supervisor module. This is especially important to end-of-row/modular switches, as the CIR applies to the aggregate traffic of all the modules CPU-bound traffic.

Applicable Hardware

The component covered in this document is applicable to all Cisco Nexus data center switches.

Interpretation of Control Plane Policing

The focus of this document is to address the most common and critical non-default class violations seen on Nexus switches.

Standard CoPP Default Profile

To understand how to interpret CoPP, the first verification must be to ensure a profile is applied and to understand if a default profile or custom profile is applied on the switch.

Note: As best practice, all Nexus switches must have CoPP enabled. If this feature is not enabled, it can cause instability for all control plane traffic as different platforms can restrict Supervisor (SUP) bound traffic. For example, if CoPP is not enabled on a Nexus 9000, traffic destined to the SUP is rate limited to 50 pps, thus the switch is made almost inoperable. CoPP is considered a requirement on Nexus 3000 and Nexus 9000 platforms.

If CoPP is not enabled, it can be re-enabled or configured on the switch by the use of the **setup** command or by the application of one of the standard default policies under the configuration option: copp profile [dense|lenient|moderate|strict]**.**

An unprotected device does not properly classify and segregate traffic into classes and thus any denial of service behavior for a specific feature or protocol is not confined to that scope and can affect the entire control plane.

Note: CoPP policies are implemented by Ternary Content-Addressable Memory (TCAM) classification redirects, and can be seen directly under **show system internal access-list input statistics module X | b CoPP** or **show hardware access-list input entries detail.**

```
N9K1# show copp status
Last Config Operation: None
Last Config Operation Timestamp: None
Last Config Operation Status: None
Policy-map attached to the control-plane: copp-system-p-policy-strict

copp-system-p-policy-strict is one of the system default profiles, in particular the strict profile.

N9K1# show running-config copp

!Command: show running-config copp
!Running configuration last done at: Tue Apr 26 16:34:10 2022
!Time: Sun May 1 16:30:57 2022

version 10.2(1) Bios:version 05.45
copp profile strict
```

Control Plane Policing Classes

CoPP classifies traffic based on the matches that correspond to the IP or MAC ACLs, Thus, it is important to understand what traffic is classified under which class.

The classes, which are platform dependent, can vary. So, it is important to understand how to verify the classes.

For example, on Nexus 9000 top-of-rack (TOR):

```
N9K1# show policy-map interface control-plane
```

```
Control Plane

Service-policy input: copp-system-p-policy-strict
...
class-map copp-system-p-class-critical (match-any)
match access-group name copp-system-p-acl-bgp
match access-group name copp-system-p-acl-rip
match access-group name copp-system-p-acl-vpc
match access-group name copp-system-p-acl-bgp6
match access-group name copp-system-p-acl-ospf
match access-group name copp-system-p-acl-rip6
match access-group name copp-system-p-acl-eigrp
match access-group name copp-system-p-acl-ospf6
match access-group name copp-system-p-acl-eigrp6
match access-group name copp-system-p-acl-auto-rp
match access-group name copp-system-p-acl-mac-l3-isis
set cos 7
police cir 36000 kbps , bc 1280000 bytes
module 1 :
transmitted 177446058 bytes;
5-minute offered rate 3 bytes/sec
conformed 27 peak-rate bytes/sec
at Sat Apr 23 04:25:27 2022

dropped 0 bytes;
5-min violate rate 0 byte/sec
violated 0 peak-rate byte/sec
...
```

In this example, the class-map copp-system-p-class-critical encompasses traffic related to routing protocols, such as Border Gateway Protocol (BGP), Open Shortest Path First (OSPF), Enhanced Interior Gateway Router Protocol (EIGRP), and include other protocols, such as vPC.

The IP or MAC ACLs name convention is mostly self-explanatory for the protocol or feature involved, with the prefix copp-system-p-acl-[protocol|feature].

To view a specific class, it can be specified directly while the **show** command is run. For example:

```
N9K-4# show policy-map interface control-plane class copp-system-p-class-management
Control Plane

Service-policy input: copp-system-p-policy-strict

class-map copp-system-p-class-management (match-any)
match access-group name copp-system-p-acl-ftp
match access-group name copp-system-p-acl-ntp
match access-group name copp-system-p-acl-ssh
match access-group name copp-system-p-acl-http
match access-group name copp-system-p-acl-ntp6
match access-group name copp-system-p-acl-sftp
match access-group name copp-system-p-acl-snmp
match access-group name copp-system-p-acl-ssh6
match access-group name copp-system-p-acl-tftp
match access-group name copp-system-p-acl-https
match access-group name copp-system-p-acl-snmp6
match access-group name copp-system-p-acl-tftp6
match access-group name copp-system-p-acl-radius
match access-group name copp-system-p-acl-tacacs
match access-group name copp-system-p-acl-telnet
match access-group name copp-system-p-acl-radius6
match access-group name copp-system-p-acl-tacacs6
match access-group name copp-system-p-acl-telnet6
set cos 2
police cir 36000 kbps , bc 512000 bytes
module 1 :
transmitted 0 bytes;
5-minute offered rate 0 bytes/sec
conformed 0 peak-rate bytes/sec

dropped 0 bytes;
5-min violate rate 0 byte/sec
violated 0 peak-rate byte/sec
```

While the CoPP default profiles are normally hidden as part of the default configuration, you can see the configuration with **show running-conf copp all**:

<#root>

```
N9K1# show running-config copp all

!Command: show running-config copp all
!Running configuration last done at: Tue Apr 26 16:34:10 2022
!Time: Sun May 1 16:41:55 2022

version 10.2(1) Bios:version 05.45
control-plane
scale-factor 1.00 module 1
class-map type control-plane match-any copp-system-p-class-critical
match access-group name

copp-system-p-acl-bgp


match access-group name copp-system-p-acl-rip
match access-group name copp-system-p-acl-vpc
match access-group name copp-system-p-acl-bgp6
match access-group name copp-system-p-acl-ospf
match access-group name copp-system-p-acl-rip6
match access-group name copp-system-p-acl-eigrp
match access-group name copp-system-p-acl-ospf6
match access-group name copp-system-p-acl-eigrp6
match access-group name copp-system-p-acl-auto-rp
match access-group name copp-system-p-acl-mac-l3-isis
(snip)
...
```

The class-map copp-system-p-class-critical, seen before, references multiple match statements that call upon system ACLs, which by default are hidden, and reference the classification that is matched upon. For example, for BGP:

<#root>

```
N9K1# show running-config aclmgr all | b

copp-system-p-acl-bgp


ip access-list

copp-system-p-acl-bgp


10 permit tcp any gt 1023 any eq bgp
20 permit tcp any eq bgp any gt 1023
(snip)
```

This means that any BGP traffic matches this class and is classified under copp-system-p-class-critical, along with all other protocols on that same class.

The Nexus 7000 uses a very similar CoPP feature structure to the Nexus 9000:

```
N77-A-Admin# show policy-map interface control-plane
Control Plane
service-policy input copp-system-p-policy-strict

class-map copp-system-p-class-critical (match-any)
match access-group name copp-system-p-acl-bgp
match access-group name copp-system-p-acl-rip
match access-group name copp-system-p-acl-vpc
match access-group name copp-system-p-acl-bgp6
match access-group name copp-system-p-acl-lisp
match access-group name copp-system-p-acl-ospf
match access-group name copp-system-p-acl-rip6
match access-group name copp-system-p-acl-rise
match access-group name copp-system-p-acl-eigrp
match access-group name copp-system-p-acl-lisp6
match access-group name copp-system-p-acl-ospf6
match access-group name copp-system-p-acl-rise6
match access-group name copp-system-p-acl-eigrp6
match access-group name copp-system-p-acl-otv-as
match access-group name copp-system-p-acl-mac-l2pt
match access-group name copp-system-p-acl-mpls-ldp
match access-group name copp-system-p-acl-mpls-rsvp
match access-group name copp-system-p-acl-mac-l3-isis
match access-group name copp-system-p-acl-mac-otv-isis
match access-group name copp-system-p-acl-mac-fabricpath-isis
match protocol mpls router-alert
set cos 7
police cir 36000 kbps bc 250 ms
conform action: transmit
violate action: drop
module 1:
conformed 300763871 bytes,
5-min offered rate 132 bytes/sec
peak rate 125 bytes/sec at Sun May 01 09:50:51 2022
violated 0 bytes,
5-min violate rate 0 bytes/sec
peak rate 0 bytes/sec
module 2:
```

```
conformed 4516900216 bytes,
5-min offered rate 1981 bytes/sec
peak rate 1421 bytes/sec at Fri Apr 29 15:40:40 2022
violated 0 bytes,
5-min violate rate 0 bytes/sec
peak rate 0 bytes/sec
module 6:
conformed 0 bytes,
5-min offered rate 0 bytes/sec
peak rate 0 bytes/sec
violated 0 bytes,
5-min violate rate 0 bytes/sec
peak rate 0 bytes/sec
```

It is important to note that on a Nexus 7000, as these are modular switches, you see the class divided by module; however, the CIR applies to the aggregate of all modules, and CoPP applies to the entire chassis. The CoPP verification and outputs can only be seen from the default or admin Virtual Device Context (VDC).

It is especially important to verify CoPP on a Nexus 7000 if control plane issues are seen, because instability on a VDC with excessive CPU-bound traffic that causes CoPP violations can impact the stability of other VDCs.

On a Nexus 5600 the classes vary. Thus, for BGP it is its own separate class:

```
N5K# show policy-map interface control-plane
Control Plane
(snip)
class-map copp-system-class-bgp (match-any)
match protocol bgp
police cir 9600 kbps , bc 4800000 bytes
conformed 1510660 bytes; action: transmit
violated 0 bytes;
(snip)
```

On a Nexus 3100, there are 3 routing protocol classes, so to verify which class BGP belongs to, cross-reference the 4 CoPP ACL that is referenced:
EIGRP is handled by its own class on the Nexus 3100.

<#root>

```
N3K-C3172# show policy-map interface control-plane
Control Plane

service-policy input: copp-system-policy

class-map copp-s-routingProto2 (match-any)
match access-group name copp-system-acl-routingproto2
police pps 1300
OutPackets 0
DropPackets 0
class-map copp-s-v6routingProto2 (match-any)
```

```
match access-group name copp-system-acl-v6routingProto2
police pps 1300
OutPackets 0
DropPackets 0
class-map copp-s-eigrp (match-any)
match access-group name copp-system-acl-eigrp
match access-group name copp-system-acl-eigrp6
police pps 200
OutPackets 0
DropPackets 0
class-map copp-s-routingProto1 (match-any)
match access-group name
```

copp-system-acl-routingproto1

```
match access-group name copp-system-acl-v6routingproto1
police pps 1000
OutPackets 0
DropPackets 0

N3K-C3172# show running-config aclmgr

!Command: show running-config aclmgr
!No configuration change since last restart
!Time: Sun May 1 18:14:16 2022

version 9.3(9) Bios:version 5.3.1
ip access-list copp-system-acl-eigrp
10 permit eigrp any 224.0.0.10/32
ipv6 access-list copp-system-acl-eigrp6
10 permit eigrp any ff02::a/128
ip access-list
```

copp-system-acl-routingproto1

10 permit tcp any gt 1024 any eq bgp

20 permit tcp any eq bgp any gt 1024

```
30 permit udp any 224.0.0.0/24 eq rip
40 permit tcp any gt 1024 any eq 639
50 permit tcp any eq 639 any gt 1024
70 permit ospf any any
80 permit ospf any 224.0.0.5/32
90 permit ospf any 224.0.0.6/32
ip access-list copp-system-acl-routingproto2
10 permit udp any 224.0.0.0/24 eq 1985
20 permit 112 any 224.0.0.0/24
ipv6 access-list copp-system-acl-v6routingProto2
10 permit udp any ff02::66/128 eq 2029
20 permit udp any ff02::fb/128 eq 5353
30 permit 112 any ff02::12/128
ipv6 access-list copp-system-acl-v6routingproto1
10 permit 89 any ff02::5/128
20 permit 89 any ff02::6/128
30 permit udp any ff02::9/128 eq 521
```

In this case, BGP is matched by the ACL copp-system-acl-routingproto1, and thus the CoPP class BGP falls into is copp-s-routingProto1.

Control Plane Policing Statistics and Counters

CoPP supports QoS statistics to track the aggregate counters of traffic that confirms or violates the committed input rate (CIR) for a particular class, for every module.

Each class-map classifies CPU bound traffic, based on the class it corresponds to and attaches a CIR for all packets that fall under that classification. As an example, the class that relates to BGP traffic is used as a reference:

On a Nexus 9000 top-of-rack (TOR) for copp-system-p-class-critical:

<#root>

```
class-map copp-system-p-class-critical (match-any)
match access-group name

copp-system-p-acl-bgp


match access-group name copp-system-p-acl-rip
match access-group name copp-system-p-acl-vpc
match access-group name copp-system-p-acl-bgp6
match access-group name copp-system-p-acl-ospf
match access-group name copp-system-p-acl-rip6
match access-group name copp-system-p-acl-eigrp
match access-group name copp-system-p-acl-ospf6
match access-group name copp-system-p-acl-eigrp6
match access-group name copp-system-p-acl-auto-rp
match access-group name copp-system-p-acl-mac-l3-isis
set cos 7
police cir 36000 kbps , bc 1280000 bytes
module 1 :
transmitted 177446058 bytes;
5-minute offered rate 3 bytes/sec
conformed 27 peak-rate bytes/sec
at Sat Apr 23 04:25:27 2022

dropped 0 bytes;
5-min violate rate 0 byte/sec
violated 0 peak-rate byte/sec
```

On the section of the class-map, after the match statements, you see the actions that relate to all traffic within the class. All traffic classified within copp-system-p-class-critical is set with a Class of Service (CoS) of 7, which is the highest priority traffic, and this class is policed with a CIR of 36000 kbps and a committed-burst-rate of 1280000 bytes.

Traffic that conforms to this policy is forwarded to the SUP to be processed and any violations are dropped.

<#root>

```
set cos 7

police cir 36000 kbps , bc 1280000 bytes
```

The next section contains the statistics that relate to the module, for top-of-rack (TOR) switches, with a single module, module 1 refers to the switch.

```
module 1 :
transmitted 177446058 bytes;
5-minute offered rate 3 bytes/sec
conformed 27 peak-rate bytes/sec
at Sat Apr 23 04:25:27 2022

dropped 0 bytes;
5-min violate rate 0 byte/sec
violated 0 peak-rate byte/sec
```

The statistics seen on the output are historical, thus this provides a snapshot of the current statistics at the time the command is run.

There are two sections to interpret here: the transmitted and dropped sections:
The transmitted datapoint tracks all packets transmitted that conform with the policy. This section is important as it provides insight into the type of traffic the supervisor processes.

The 5-minute offered rate value provides insight into the current rate.
The conformed peak rate and date, provides a snap of the highest peak-rate per seconds that still conformed within the policy and the time it occurred.
If a new peak is seen, then it replaces this value and date.

The most important portion of the statistics is the dropped datapoint. Just like the transmitted statistics, the dropped section tracks the cumulative bytes dropped due to violations to the police rate. It also provides the violation rate for the last 5 minutes, the violated peak, and if there is a peak, the timestamp of that peak violation. And again, if a new peak is seen, then it replaces this value and date. On other platforms, the outputs vary, but the logic is very similar.

Nexus 7000 uses an identical structure and the verification is the same, though some classes vary slightly on the ACLs referenced:

<#root>

```
class-map

copp-system-p-class-critical

 (match-any)
match access-group name

copp-system-p-acl-bgp


match access-group name copp-system-p-acl-rip
match access-group name copp-system-p-acl-vpc
match access-group name copp-system-p-acl-bgp6
match access-group name copp-system-p-acl-lisp
match access-group name copp-system-p-acl-ospf
match access-group name copp-system-p-acl-rip6
match access-group name copp-system-p-acl-rise
match access-group name copp-system-p-acl-eigrp
match access-group name copp-system-p-acl-lisp6
match access-group name copp-system-p-acl-ospf6
match access-group name copp-system-p-acl-rise6
match access-group name copp-system-p-acl-eigrp6
match access-group name copp-system-p-acl-otv-as
match access-group name copp-system-p-acl-mac-l2pt
match access-group name copp-system-p-acl-mpls-ldp
match access-group name copp-system-p-acl-mpls-rsvp
match access-group name copp-system-p-acl-mac-l3-isis
match access-group name copp-system-p-acl-mac-otv-isis
match access-group name copp-system-p-acl-mac-fabricpath-isis
match protocol mpls router-alert

set cos 7



police cir 36000 kbps bc 250 ms



conform action: transmit


violate action: drop
module 1:
conformed 300763871 bytes,
5-min offered rate 132 bytes/sec
peak rate 125 bytes/sec at Sun May 01 09:50:51 2022
violated 0 bytes,
5-min violate rate 0 bytes/sec
peak rate 0 bytes/sec
module 2:
conformed 4516900216 bytes,
5-min offered rate 1981 bytes/sec
peak rate 1421 bytes/sec at Fri Apr 29 15:40:40 2022
violated 0 bytes,
5-min violate rate 0 bytes/sec
peak rate 0 bytes/sec
module 6:
conformed 0 bytes,
5-min offered rate 0 bytes/sec
peak rate 0 bytes/sec
violated 0 bytes,
5-min violate rate 0 bytes/sec
peak rate 0 bytes/sec
```

On a Nexus 5600:

<#root>

```
class-map copp-system-class-bgp

 (match-any)

match protocol bgp
```

```
police cir 9600 kbps , bc 4800000 bytes
conformed 1510660 bytes; action: transmit
violated 0 bytes;
```

While it does not provide information on rate or peaks, it still provides the aggregate bytes conformed and violated.

On a Nexus 3100, the control plane output shows, OutPackets and DropPackets.

```
class-map copp-s-routingProto1 (match-any)
match access-group name copp-system-acl-routingproto1
match access-group name copp-system-acl-v6routingproto1
police pps 1000
OutPackets 8732060
DropPackets 0
```

OutPackets refer to conformed packets, while DropPackets refer to violations to the CIR. In this scenario, you see no drops on the associated class.

On a Nexus 3500, the output shows HW and SW Matched Packets:

```
class-map copp-s-routingProto1 (match-any)
match access-group name copp-system-acl-routingproto1
police pps 900
HW Matched Packets 471425
SW Matched Packets 471425
```

The HW Matched Packets refer to the packets that are matched in HW by the ACL. The SW matched packets are the ones that comply with the Policy. Any differences between the HW and SW matched packets implies a violation.

In this case, there are no drops seen on routing protocol-1 class packets (which includes BGP), as the values match.

Check for Active Drop Violations

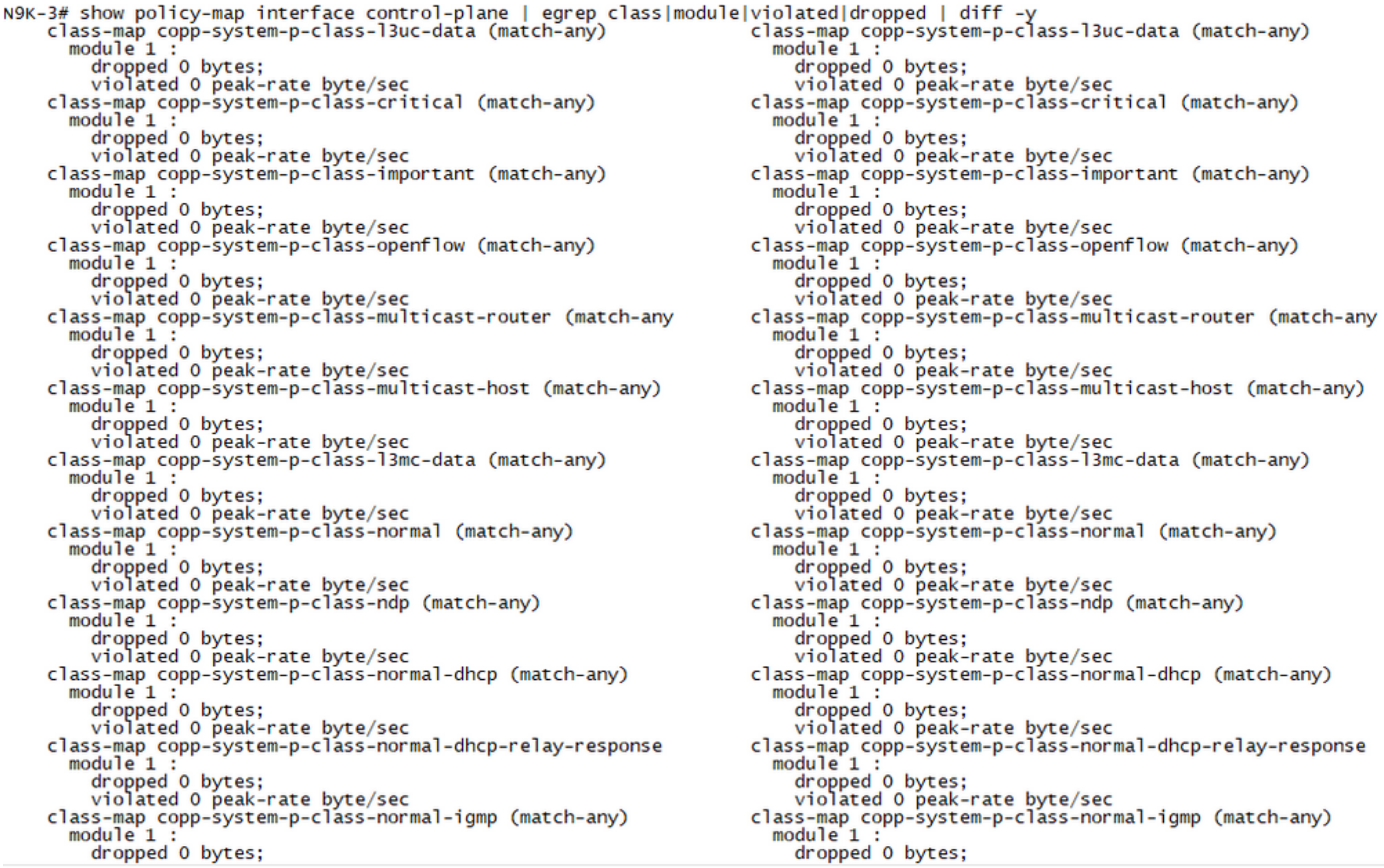
Given that the control plane policing statistics are historical, it is important to determine if active violations are on the increase. The standard way to perform this task is to compare two full outputs and verify any differences.

This task can be performed manually, or the Nexus switches provide the diff tool that can assist to compare the outputs.

While the entire output can be compared, it is not required because the focus is only on the dropped statistics. Thus, the CoPP output can be filtered to focus only on the violations.

The command is: show policy-map interface control-plane | egrep class|module|violated|dropped | diff -y

Note: The command must be run twice for the diff to be able to compare the current to the previous output.



```
N9K-3# show policy-map interface control-plane | egrep class|module|violated|dropped | diff -y
    class-map copp-system-p-class-l3uc-data (match-any)             class-map copp-system-p-class-l3uc-data (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-critical (match-any)              class-map copp-system-p-class-critical (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-important (match-any)             class-map copp-system-p-class-important (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-openflow (match-any)              class-map copp-system-p-class-openflow (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-multicast-router (match-any       class-map copp-system-p-class-multicast-router (match-any
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-multicast-host (match-any)        class-map copp-system-p-class-multicast-host (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-l3mc-data (match-any)             class-map copp-system-p-class-l3mc-data (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-normal (match-any)                class-map copp-system-p-class-normal (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-ndp (match-any)                   class-map copp-system-p-class-ndp (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-normal-dhcp (match-any)           class-map copp-system-p-class-normal-dhcp (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-normal-dhcp-relay-response        class-map copp-system-p-class-normal-dhcp-relay-response
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
        violated 0 peak-rate byte/sec                                   violated 0 peak-rate byte/sec
    class-map copp-system-p-class-normal-igmp (match-any)           class-map copp-system-p-class-normal-igmp (match-any)
      module 1 :                                                      module 1 :
        dropped 0 bytes;                                                dropped 0 bytes;
```

The previous command allows you to see the delta between two classes and find violation increases.

Note: As the CoPP statistics are historical, another recommendation is to clear the statistics after the command is run, to verify if there are active increases. To clear the CoPP statistics, run the command: **clear copp statistics.**

Types of CoPP Drops

CoPP is a simple policing structure, as any CPU bound traffic that violates the CIR is dropped. The implications nonetheless vary significantly dependent on the type of drops.

While the logic is the same, it is not the same to drop traffic destined to copp-system-p-class-critical.

```
class-map copp-system-p-class-critical (match-any)
match access-group name copp-system-p-acl-bgp
match access-group name copp-system-p-acl-rip
match access-group name copp-system-p-acl-vpc
match access-group name copp-system-p-acl-bgp6
match access-group name copp-system-p-acl-ospf
match access-group name copp-system-p-acl-rip6
match access-group name copp-system-p-acl-eigrp
match access-group name copp-system-p-acl-ospf6
match access-group name copp-system-p-acl-eigrp6
match access-group name copp-system-p-acl-auto-rp
match access-group name copp-system-p-acl-mac-l3-isis
set cos 7
police cir 36000 kbps , bc 1280000 bytes
```

Compared to drop traffic destined to class-map copp-system-p-class-monitoring.

```
class-map copp-system-p-class-monitoring (match-any)
match access-group name copp-system-p-acl-icmp
match access-group name copp-system-p-acl-icmp6
match access-group name copp-system-p-acl-traceroute
set cos 1
police cir 360 kbps , bc 128000 bytes
```

The first one deals with mostly routing protocols, the second one deals with Internet Control Message Protocol (ICMP) which has one of the lowest priorities and CIR. The difference on CIR is one hundred-fold. Therefore, it is important to understand the classes, impacts, common checks/verifications, and recommendations.

CoPP Classes

Class Monitoring - copp-system-p-class-monitoring

This class encompasses ICMP for IPv4, and IPv6, and traceroute of traffic directed to the switch in question.

```
class-map copp-system-p-class-monitoring (match-any)
match access-group name copp-system-p-acl-icmp
match access-group name copp-system-p-acl-icmp6
match access-group name copp-system-p-acl-traceroute
set cos 1
police cir 360 kbps , bc 128000 bytes
```

Impact

A common misconception when packet loss or latency is troubleshot is to ping the switch through its in-band ports, which are rate-limited by CoPP. As CoPP heavily polices ICMP, even with a low traffic or congestion, packet loss can be seen by a ping to in-band interfaces directly if they violate the CIR.

For example, by a ping to directly connected interfaces on routed ports, with a packet payload of 500, drops can be seen periodically.

<#root>

```
N9K-3# ping 192.168.1.1 count 1000 packet-size 500
...
--- 192.168.1.1 ping statistics ---
1000 packets transmitted, 995 packets received,

0.50% packet loss


round-trip min/avg/max = 0.597/0.693/2.056 ms
```

On the Nexus, where the ICMP packets were destined, you see that CoPP dropped them as the violation was detected and the CPU was protected:

<#root>

```
N9K-4# show policy-map interface control-plane class copp-system-p-class-monitoring
Control Plane

Service-policy input: copp-system-p-policy-strict

class-map copp-system-p-class-monitoring (match-any)
match access-group name copp-system-p-acl-icmp
match access-group name copp-system-p-acl-icmp6
match access-group name copp-system-p-acl-traceroute
set cos 1
police cir 360 kbps , bc 128000 bytes
module 1 :
transmitted 750902 bytes;
5-minute offered rate 13606 bytes/sec
conformed 13606 peak-rate bytes/sec
at Sun May 01 22:49:24 2022


dropped 2950 bytes;



5-min violate rate 53 byte/sec



violated 53 peak-rate byte/sec at Sun May 01 22:49:24 2022
```

To troubleshoot latency or packet loss, it is recommended to use hosts reachable through the switch by the data plane, not destined to the switch itself which would be control plane traffic. Data plane traffic is forwarded/routed at the hardware level without SUP intervention and thus not policed by CoPP, and typically experience no drops.

Recommendations

- Send a ping across the switch through the data plane, not to the switch, to verify false positive results for packet loss.
- Limit Network Monitoring System (NMS) or tools that aggressively use ICMP the switch to avoid a burst through the committed input rate for the class. Remember that CoPP applies to all aggregate traffic that falls into the class.

Class Management - copp-system-p-class-management

As seen here, this class encompasses different management protocols that can be used for communication (SSH, Telnet), transfers (SCP, FTP, HTTP, SFTP, TFTP), clock (NTP), AAA (Radius/TACACS) and monitoring (SNMP), for IPv4 and IPv6 communications.

```
class-map copp-system-p-class-management (match-any)
match access-group name copp-system-p-acl-ftp
match access-group name copp-system-p-acl-ntp
match access-group name copp-system-p-acl-ssh
match access-group name copp-system-p-acl-http
match access-group name copp-system-p-acl-ntp6
match access-group name copp-system-p-acl-sftp
match access-group name copp-system-p-acl-snmp
match access-group name copp-system-p-acl-ssh6
match access-group name copp-system-p-acl-tftp
match access-group name copp-system-p-acl-https
match access-group name copp-system-p-acl-snmp6
match access-group name copp-system-p-acl-tftp6
match access-group name copp-system-p-acl-radius
match access-group name copp-system-p-acl-tacacs
match access-group name copp-system-p-acl-telnet
match access-group name copp-system-p-acl-radius6
match access-group name copp-system-p-acl-tacacs6
match access-group name copp-system-p-acl-telnet6
set cos 2
police cir 36000 kbps , bc 512000 bytes
```

Impact

The most common behaviors or drops associated with this class include:

- Perceived CLI slowness when connected by SSH/Telnet. If there are active drops on the class, then communication sessions can be slow and suffer from drops.
- Transfer files with FTP, SCP, SFTP, TFTP protocols on the switch. The most common behavior seen is an attempt to transfer system/kickstart boot images by in-band management ports. This can lead to higher transfer times and closed/terminated transmission sessions determined by the aggregate bandwidth for the class.
- NTP synchronization issues, this class is also important because it mitigates rogue NTP agents or attacks.
- AAA Radius and TACACS services also fall in this class. If impact is perceived on this class, it can affect authorization and authentication services on the switch for user-accounts, which can also contribute to delay on the CLI commands.
- SNMP is also policed under this class. The most common behavior seen due to drops due to the SNMP class are on NMS servers, which perform walks, bulk collections, or network scans. When periodic instability occurs, usually it is correlated to the NMS collection schedule.

Recommendations

- If CLI slowness is perceived, along with drops in this class, use console access, or management out-of-band access (mgmt0).
- If system images must be uploaded to the switch, use either the out-of-band management port (mgmt0) or use the USB ports for the fastest transfer.
- If NTP packets are lost, check show ntp peer-status, and verify the reachability column, no drops do translate to 377.
- If issues are seen with AAA services, use local-only users to troubleshoot, until behavior is mitigated.

- Mitigation for SNMP issues include less aggressive behavior, targeted collection, or minimization of network scanners. Examine periodic times from scanners to events seen at the CPU level.

Class L3 Unicast Data - copp-system-p-class-l3uc-data

This class deals specifically with glean packets. This type of packet is also handled by the Hardware Rate Limiter (HWRL).

If the Address Resolution Protocol (ARP) request for the next hop is not resolved when incoming IP packets are forwarded in a line card, the line card forwards the packets to the supervisor module.

The supervisor resolves the MAC address for the next hop and programs the hardware.

```
class-map copp-system-p-class-l3uc-data (match-any)
match exception glean
set cos 1
```

This normally occurs when static routes are used and the next hop is unreachable or unresolved.

When an ARP request is sent, the software adds a /32 drop adjacency in the hardware to prevent the packets to the same next-hop IP address to be forwarded to the supervisor. When the ARP is resolved, the hardware entry is updated with the correct MAC address. If the ARP entry is not resolved before a timeout period, the entry is removed from the hardware.

Note: CoPP and HWRL work in tandem to ensure the CPU is protected. While they appear to perform similar functions, HWRL occurs first. The implementation is based on where the specific feature is implemented on the forwarding engines on the ASIC. This serial approach allows granularity and multilayer protections that rate all CPU bound packets.

The HWRL is performed per instance/forwarding engine on the module and can be viewed with the command **show hardware rate-limiter**. HWRL is outside of the scope of this technical document.

<#root>

```
show hardware rate-limiter

Units for Config: kilo bits per second
Allowed, Dropped & Total: aggregated bytes since last clear counters


Module: 1
R-L Class Config Allowed Dropped Total
+----------------+----------+-------------------+-------------------+-------------------+

L3 glean 100 0 0 0


L3 mcast loc-grp 3000 0 0 0
access-list-log 100 0 0 0
bfd 10000 0 0 0
fex 12000 0 0 0
```

```
span 50 0 0 0
sflow 40000 0 0 0
vxlan-oam 1000 0 0 0
100M-ethports 10000 0 0 0
span-egress disabled 0 0 0
dot1x 3000 0 0 0
mpls-oam 300 0 0 0
netflow 120000 0 0 0
ucs-mgmt 12000 0 0 0
```

Impact

- Data plane traffic is punted to the supervisor as a violation, as it cannot be processed in hardware, and thus creates pressure on the CPU.

Recommendations

- The common resolution for this matter to minimize the glean drops is to ensure the next hop is reachable, and to enable glean-throttling by the configuration command: **hardware ip glean throttle**.

On Nexus 7000 8.4(2), it also introduced bloom filter support for glean adjacencies for M3 and F4 modules. Refer to: Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide

Review any static route configurations that use unreachable next-hop addresses, or use dynamic routing protocols that would remove such routes from the RIB dynamically.

Class Critical - class-map copp-system-p-class-critical

This class references the most critical control plane protocols from a L3 perspective, which include routing protocols for IPv4 and IPv6, (RIP, OSPF, EIGRP, BGP), auto-RP, virtual port-channel (vPC), and l2pt and IS-IS.

```
class-map copp-system-p-class-critical (match-any)
match access-group name copp-system-p-acl-bgp
match access-group name copp-system-p-acl-rip
match access-group name copp-system-p-acl-vpc
match access-group name copp-system-p-acl-bgp6
match access-group name copp-system-p-acl-ospf
match access-group name copp-system-p-acl-rip6
match access-group name copp-system-p-acl-eigrp
match access-group name copp-system-p-acl-ospf6
match access-group name copp-system-p-acl-eigrp6
match access-group name copp-system-p-acl-auto-rp
match access-group name copp-system-p-acl-mac-l2pt
match access-group name copp-system-p-acl-mac-l3-isis
set cos 7
police cir 36000 kbps , bc 1280000 bytes
```

Impact

Drops on copp-system-p-class-critical transfer instability to routing protocols, which can include adjacencies

dropped or convergence failures, or update/NLRI propagation.
The most common policy drops on this class can relate to rogue devices on the network that act abnormally (due to misconfiguration or failure) or scalability.

Recommendations

- If there are no anomalies detected, like a rogue device or L2 instability that causes continuous reconvergence of upper layer protocols, then a custom configuration of CoPP or a more lenient class can be required to accommodate the scale.
- Refer to the CoPP configuration guide for how to configure a custom CoPP profile from a default profile that currently exists.
  Copying the CoPP Best Practice Policy

Class Important - copp-system-p-class-important

This class relates to the first-hop redundancy protocols (FHRP), which includes HSRP, VRRP, and also LLDP

```
class-map copp-system-p-class-important (match-any)
match access-group name copp-system-p-acl-hsrp
match access-group name copp-system-p-acl-vrrp
match access-group name copp-system-p-acl-hsrp6
match access-group name copp-system-p-acl-vrrp6
match access-group name copp-system-p-acl-mac-lldp
set cos 6
police cir 2500 kbps , bc 1280000 bytes
```

Impact

The most common behavior seen here that leads to drops, are issues with Layer 2 instability, which leads to devices that transition into active state (split brain) scenarios, aggressive timers, misconfigurations, or scalability.

Recommendations:

- Ensure for FHRP that groups are properly configured and the roles are either active/standby, or primary/secondary, and are properly negotiated, and there are no flaps on the state.
- Check for convergence issues at L2 or issues with multicast propagation for the L2 domain.

Class L2 Unpoliced - copp-system-p-class-l2-unpoliced

The L2 unpoliced class refers to all critical Layer 2 protocols that are the groundwork for all upper layer protocols and thus are considered almost unpoliced with the highest CIR and priority.

Effectively, this class handles, Spanning-Tree Protocol (STP), Link Aggregation Control Protocol (LACP), Cisco Fabric Service over Ethernet (CFSoE)

```
class-map copp-system-p-class-l2-unpoliced (match-any)
```

```
match access-group name copp-system-p-acl-mac-stp
match access-group name copp-system-p-acl-mac-lacp
match access-group name copp-system-p-acl-mac-cfsoe
match access-group name copp-system-p-acl-mac-sdp-srp
match access-group name copp-system-p-acl-mac-l2-tunnel
match access-group name copp-system-p-acl-mac-cdp-udld-vtp
set cos 7
police cir 50 mbps , bc 8192000 bytes
```

This class has a police CIR of 50 Mbps, the highest amongst all classes, along with the highest burst rate absorption.

Impact

Drops on this class can lead to global instability, as all upper layer protocols and communications on data, control, and management planes rely on an underlying Layer 2 stability.

Issues with STP violations can cause TCNs and STP convergence issues, which includes STP disputes, MAC flushes, moves, and learning disabled behaviors, that causes reachability issues and can cause traffic loops that destabilize the network.

This class also references LACP, and thus handles all EtherType packets associated with 0x8809, which include all LACPDUs used to maintain the state of the port-channel bonds. Instability on this class, can cause the port-channels to timeout if the LACPDUs are dropped.

Cisco Fabric Service over Ethernet (CSFoE) falls within this class and is used to communicate critical application control states between Nexus switches and therefore is imperative for stability.

The same applies to other protocols within this class, which includes CDP, UDLD, and VTP.

Recommendations

- The most common behavior relates to L2 Ethernet instability. Ensure STP is properly designed in a deterministic way with the relevant feature enhancements in play to minimize the impact of reconvergence or rogue devices in the network. Make sure the proper STP port type is configured for all end-host devices that do not participate on the L2 extension are configured as edge/edge trunk ports to minimize TCNs.
- Use STP enhancements, such as BPDUguard, Loopguard, BPDUfilter, and RootGuard where appropriate to limit the scope of a failure, or issues with misconfiguration or rogue devices on the network.
- Refer to: Cisco Nexus 9000 NX-OS Layer 2 Switching Configuration Guide, Release 10.2(x)
- Check for MAC move behaviors that can lead to disablement of MAC learning and flushes. Refer to:Nexus 9000 Mac move troubleshooting and preventive methods

Class Multicast Router - class-map copp-system-p-class-multicast-router

This class refers to control plane Protocol Independent Multicast (PIM) packets used for the establishment and control of routed multicast-shared trees through all PIM enabled devices in the data plane path, and includes First-Hop Router (FHR), Last-Hop Router (LHR), Intermediate-Hop Routers (IHR), and Rendezvous Points (RPs). Packets classified within this class include PIM registration for sources, PIM joins for receivers for both IPv4 and IPv6, in general any traffic destined for PIM (224.0.0.13), and Multicast Source Discovery Protocol (MSDP). Be aware that there are several additional classes, which deal with very specific portions of multicast or RP functionality that are handled by different classes.

```
class-map copp-system-p-class-multicast-router (match-any)
match access-group name copp-system-p-acl-pim
match access-group name copp-system-p-acl-msdp
match access-group name copp-system-p-acl-pim6
match access-group name copp-system-p-acl-pim-reg
match access-group name copp-system-p-acl-pim6-reg
match access-group name copp-system-p-acl-pim-mdt-join
match exception mvpn
set cos 6
police cir 2600 kbps , bc 128000 bytes
```

Impact

The main impact on drops that relate to this class are associated with issues that communicate to multicast sources by PIM registration toward the RPs or PIM joins not properly processed, which would destabilize the shared or shortest path trees toward the sources of the multicast stream or to the RPs. Behavior can include outgoing interface list (OIL) not properly populated due to absent joins, or (S, G), or (*, G) not seen consistently across the environment. Issues can also arise between multicast routing domains that rely on MSDP for interconnection.

Recommendations

- The most common behavior for PIM control-related issues refer to scale issues, or rogue behaviors. One of the most common behaviors is seen due to the implementation on UPnP, which can also cause memory exhaustion issues. This can be addressed by filters and reduced scope of the rogue devices. For details on how the mitigate and filter multicast control packets that depend on the network role of the device, refer to: Configure Multicast Filtering on Nexus 7K/N9K - Cisco

Class Multicast Host - copp-system-p-class-multicast-host

This class refers to Multicast Listener Discovery (MLD), specifically MLD query, report, reduction, and MLDv2 packet types. MLD is an IPv6 protocol a host uses to request multicast data for a particular group. With the information obtained through MLD, the software maintains a list of multicast group or channel memberships on a per-interface basis. The devices that receive MLD packets send the multicast data they receive for requested groups or channels out the network segment of the known receivers. MLDv1 is derived from IGMPv2, and MLDv2 is derived from IGMPv3. IGMP uses IP Protocol 2 message types, while MLD uses IP Protocol 58 message types, which is a subset of the ICMPv6 messages.

```
class-map copp-system-p-class-multicast-host (match-any)
match access-group name copp-system-p-acl-mld
set cos 1
police cir 1000 kbps , bc 128000 bytes
```

Impact

Drops on this class translate to issues on link-local IPv6 multicast communications, which can cause listener reports from receivers or responses to general queries to be dropped, which prevents discovery of multicast groups the hosts want to receive. This can impact the snooping mechanism and not properly forward traffic out through expected interfaces that requested the traffic.

Recommendations

- As MLD traffic is significant on a link-local level for IPv6, if drops are seen on this class, the most common behavior causes relate to scale, L2 instability, or rogue devices.

Class Layer 3 Multicast Data - copp-system-p-class-l3mc-data and Class Layer 3 Multicast IPv6 Data - copp-system-p-class-l3mcv6-data

These classes refer to traffic that matches a multicast exception redirection toward the SUP. In this case, there are two conditions that are handled by these classes. The first is Reverse-Path Forwarding (RPF) failure and the second is Destination Miss. Destination Miss refers to multicast packets where the lookup in hardware for the Layer 3 multicast forwarding table fails, and thus the data packet is punted to the CPU. These packets are sometimes used to trigger/install the multicast control plane and add the hardware forwarding tables entries, based on the data plane traffic. Data plane multicast packets that violate the RPF would also match this exception and be classified as a violation.

```
class-map copp-system-p-class-l3mc-data (match-any)
match exception multicast rpf-failure
match exception multicast dest-miss
set cos 1
police cir 2400 kbps , bc 32000 bytes

class-map copp-system-p-class-l3mcv6-data (match-any)
match exception multicast ipv6-rpf-failure
match exception multicast ipv6-dest-miss
set cos 1
police cir 2400 kbps , bc 32000 bytes
```

Impact

RPF failures and Destination misses imply a design or configuration issue related to how traffic flows through the multicast router. Destination misses are common at state creation, drops can lead to programming and creation of (*, G), (S, G) failures.

Recommendations

- Perform changes to the foundational unicast RIB design, or add static mroute to steer traffic through a particular interface, in the case of RPF failures.
- Refer to Router Does Not Forward Multicast Packets to Host Due to RPF Failure

Class IGMP - copp-system-p-class-igmp

This class refers to all IGMP messages, for all versions that are used to request multicast data for a particular group, and used by the IGMP snooping functionality to maintain the groups and relevant outgoing interface list (OIL) that forwards the traffic through to the interested receivers at Layer 2. The IGMP messages are locally significant because they do not traverse a Layer 3 boundary, as their time to live (TTL) must be 1, as documented under RFC2236 (Internet Group Management Protocol, Version 2). The IGMP packets handled by this class include all membership queries (general or source/group specific), along with the membership and leave reports from the receivers.

```
class-map copp-system-p-class-normal-igmp (match-any)
match access-group name copp-system-p-acl-igmp
set cos 3
police cir 3000 kbps , bc 64000 bytes
```

Impact

Drops on this class would translate to issues at all levels of a multicast communication between source and receiver, dependent on the type of IGMP message dropped due to the violation. If membership reports from receivers are lost, then the router is not aware of devices interested on the traffic and thus it does not include the interface/VLAN on its relevant outgoing interface list. If this device is also the querier or designated router, it does not trigger the relevant PIM join messages toward the RP if the source is beyond the local Layer 2 domain, thus it never establishes the data plane across the multicast tree all the way to the receiver or RP. If the leave report is lost, the receiver can continue to receive unwanted traffic. This can also affect all relevant IGMP queries triggered by the querier and communication between the multicast routers in a domain.

Recommendations

- The most common behaviors associated with IGMP drops relate to L2 instability, issues with timers, or scale.

Class Normal - copp-system-p-class-normalcopp-system-p-class-normal

This class refers to traffic that matches standard ARP traffic, and also includes traffic associated with 802.1X, used for port-based network access control. This is one of the most common classes which encounters violations as ARP requests, Gratuitous ARP, Reverse ARP packets are broadcast and propagate through the entire Layer 2 domain. It is important to remember that ARP packets are not IP packets, these packets do not contain a L3 header, and so the decision is made purely on the scope of the L2 headers. If a router is configured with an IP interface associated with that subnet, such as an Switch Virtual Interface (SVI), the router punts the ARP packets to the SUP to be processed, as they are destined to the hardware broadcast address. Any broadcast storm, Layer 2 loop (due to STP or flaps), or a rouge device in the network can lead to an ARP storm which causes violations to increase significantly.

```
class-map copp-system-p-class-normal (match-any)
match access-group name copp-system-p-acl-mac-dot1x
match protocol arp
set cos 1
police cir 1400 kbps , bc 32000 bytes
```

Impact

The impact of violations in this class depends heavily on the duration of the events and the role of the switch on the environment. Drops in this class imply that ARP packets are currently discarded and thus not processed by the SUP engine that can lead to two main behaviors caused by incomplete ARP resolutions.

From the perspective of the end host, devices in the network are not able to resolve or complete the address resolution with the switch. If this device acts as the default gateway for the segment, it can lead to devices unable to resolve their gateway and thus unable to route outside of their L2 Ethernet segment (VLAN). Devices can still communicate on the local segment if they can complete the ARP resolution for other end hosts on the local segment.

From the perspective of the switch, if the storm and violations are prevalent, it can also lead to the switch not able to complete the process for ARP request it generated. These requests are normally generated for next-hop or directly connected subnet resolutions. While the ARP replies are unicast in nature, as they are addressed to the MAC owned by the switch, they are classified under this same class, as they are still ARP packets. This translates into reachability issues because the switch cannot properly process traffic if the next hop is not resolved, and can lead to issues with Layer 2 header rewrite, if the adjacency manager does not have an entry for the host.

The impact also depends on the scope of the foundational issue that triggered the ARP violation. For example, in a broadcast storm, hosts and the switch continue to ARP to try to resolve the adjacency, which can lead to additional broadcast traffic on the network, and as ARP packets are Layer 2, there is no Layer 3 time to live (TTL) to break a L2 loop and thus they continue to loop, and exponentially grow through the network until the loop is broken.

Recommendations

- Resolve any foundational L2 instability that can cause ARP storms on the environment, such as STP, flaps, or rogue devices. Break those loops as required, by any desired method to open the link path.
- Storm-control can also be used to mitigate an ARP storm. If storm control is not enabled, verify counter statistics on interfaces to verify the percentage of broadcast traffic seen on the interfaces in relation to the total traffic that passes through the interface.
- If there is no storm, but constant drops are still seen on the environment, verify SUP traffic to identify any rogue devices, which constantly send ARP packets on the network, that can affect legitimate traffic.
- Increases that can be seen depend on the number of hosts on the network and role of the switch on the environment, the ARP is designed to retry, resolve, and refresh entries and thus it is expected to see ARP traffic at all times. If only sporadic drops are seen, they can be transient due to the network load and no impact is perceived. But it is important to monitor and know the network to properly identify and differentiate an expected from an abnormal situation.

Class NDP - copp-system-p-acl-ndp

This class refers to traffic associated with IPv6 neighbor discovery/advertisement and router solicitation and advertisement packets that use ICMP messages to determine local Link-layer addresses of neighbors, and it is used for reachability and track of neighbor devices.

```
class-map copp-system-p-class-ndp (match-any)
match access-group name copp-system-p-acl-ndp
set cos 6
police cir 1400 kbps , bc 32000 bytes
```

Impact

Violations on this class can impede IPv6 communication between neighbor devices, as these packets are used to facilitate the dynamic discovery or Link-layer/local information between hosts and routers on the local link. A break of this communication can also cause issues with reachability beyond or through the associated local link. If there are communication issues between IPv6 neighbors, ensure there are no drops on this class.

Recommendations

- Examine any abnormal ICMP behaviors from neighbor devices, particularly those that relate to the neighbor discovery and/or router discovery.
- Ensure all expected timer and interval values for the periodic messages are consistent across the environment, and are honored. For example, for router advertisement messages (RA messages).

Class Normal DHCP - copp-system-p-class-normal-dhcp

This class refers to traffic associated with the Bootstrap Protocol (BOOTP client/server), commonly known as Dynamic Host Control Protocol (DHCP) packets on the same local Ethernet segment for both IPv4 and IPv6. This specifically relates only to the traffic communication that originates from any bootp client or destined to any BOOTP servers, through the entire discovery, offer, request, and acknowledge (DORA) packet exchange, and also includes DHCPv6 client/server transaction through UDP ports 546/547.

```
class-map copp-system-p-class-normal-dhcp (match-any)
match access-group name copp-system-p-acl-dhcp
match access-group name copp-system-p-acl-dhcp6
set cos 1
police cir 1300 kbps , bc 32000 bytes
```

Impact

Violations on this class can lead to end hosts unable to properly acquire an IP from the DHCP server, and thus fall back to their automatic private IP address (APIPA) range, 169.254.0.0/16. Such violations can occur in environments where devices try to boot simultaneously and thus go beyond the CIR associated with the class.

Recommendations

- Verify with captures, on hosts and DHCP server side the entire DORA transaction is seen. If the switch is part of this communication, then it is also important to verify the packets processed or punted to the CPU, and verify statistics on switch: **show ip dhcp global statistics** and redirections: **show system internal access-list sup-redirect-stats module 1 | grep -i dhcp**.

Class Normal DHCP Relay Response - copp-system-p-class-normal-dhcp-relay-response

This class refers to traffic associated the DHCP relay functionality for both IPv4 and IPv6, directed to the configured DHCP servers configured under the relay. This relates specifically only to the traffic communication that originates from any BOOTP server or destined to any BOOTP clients through the entire DORA packet exchange, and also includes DHCPv6 client/server transaction through UDP ports 546/547.

```
class-map copp-system-p-class-normal-dhcp-relay-response (match-any)
match access-group name copp-system-p-acl-dhcp-relay-response
match access-group name copp-system-p-acl-dhcp6-relay-response
set cos 1
police cir 1500 kbps , bc 64000 bytes
```

Impact

Violations for this class carry the same impact as the violations for the class copp-system-p-class-normal-dhcp, because they are both parts of the same transaction. This class focuses primarily on response communications from the relay agent servers. The Nexus does not act as the DHCP server, it is designed only to act as a relay agent.

Recommendations

- The same recommendations as class normal DHCP apply here. As the function of the Nexus is only to act as a relay agent, on the SUP you expect to see the entire transaction between the host and the switch acting as relay, and the switch and the servers configure.
- Ensure there are no rogue devices, such as unexpected DHCP servers on the network that respond to the scope, or devices stuck in a loop that flood the network with DHCP Discover packets. Additional checks can be performed by the commands: show ip dhcp relay and **show ip dhcp relay statistics.**

Class NAT Flow - copp-system-p-class-nat-flow

This class refers to software switch NAT flow traffic. When a new dynamic translation is created, the flow is software forwarded until the translation is programmed in hardware, and then it is policed by CoPP to limit the traffic punted to the supervisor while the entry is installed in hardware.

```
class-map copp-system-p-class-nat-flow (match-any)
match exception nat-flow
set cos 7
police cir 800 kbps , bc 64000 bytes
```

Impact

Drops on this class typically occur when a high rate of new dynamic translations and flows are installed in hardware. The impact relates to software switched packets that are discarded and not delivered to the end host, which can lead to loss and retransmissions. Once the entry is installed in hardware, no further traffic is punted to the supervisor.

Recommendations

- Verify guidelines and limitations of dynamic NAT on the relevant platform. There are known limitations that are documented on platforms, such as the 3548 in which the translation can take a few seconds. Refer to: Restrictions for Dynamic NAT

Class Exception - copp-system-p-class-exception

This class refers to exception packets associated with IP option and IP ICMP unreachable packets. If a destination address is not present on the forwarding information base (FIB) and results in a miss, the SUP sends an ICMP unreachable packet back to the sender. Packets with IP options enabled also fall within this class., Refer to the IANA document, for details on IP options: IP Option Numbers

```
 class-map copp-system-p-class-exception (match-any)
match exception ip option
match exception ip icmp unreachable
match exception ipv6 option
match exception ipv6 icmp unreachable
set cos 1
police cir 150 kbps , bc 32000 bytes
```

Impact

This class is heavily policed, and drops on this class are not indicative of a failure but rather of a protection mechanism to limit the scope of ICMP unreachables and IP options packets.

Recommendations

- Verify if there are any flows of traffic seen or punted to the CPU for destinations not on the FIB.

Class Redirect - copp-system-p-class-redirect

This class refers to traffic associated with Precision Time Protocol (PTP), used for time synchronization. This includes multicast traffic for the reserved range 224.0.1.129/32, unicast traffic on UDP port 319/320 and Ethetype 0X88F7.

```
class-map copp-system-p-class-redirect (match-any)
match access-group name copp-system-p-acl-ptp
match access-group name copp-system-p-acl-ptp-l2
match access-group name copp-system-p-acl-ptp-uc
set cos 1
police cir 280 kbps , bc 32000 bytes
```

Impact

Drops on this class can lead to issues on devices that have not properly synched or have not established the proper hierarchy.

Recommendations

- Ensure stability of clocks, and that they are configured correctly. Make sure the PTP device is configured for multicast or unicast PTP mode, but not both at the same time. This is also documented under the guidelines and limitation, and can push the traffic beyond the committed input rate.
- Review the design and configuration of the boundary clock and all PTP devices in the environment. Ensure all guidelines and limitations are followed per platform because they vary.

Class OpenFlow - copp-system-p-class-openflow

This class refers to traffic associated with OpenFlow agent operations and the corresponding TCP connection between the controller and the agent.

```
class-map copp-system-p-class-openflow (match-any)
match access-group name copp-system-p-acl-openflow
set cos 5
police cir 1000 kbps , bc 32000 bytes
```

Impact

Drops on this class can lead to issues on agents that do not properly receive and process the instructions from the controller to manage the forwarding plane of the network

Recommendations

- Ensure no duplicate traffic is seen on the network, or any device that hinders the communication between the controller and agents.
- Verify the L2 network has no instability (STP or loops).

Troubleshoot CoPP Drops

The first steps to troubleshoot CoPP violations are to determine:

- Impact and scope of the issue.
- Understand the traffic flow through the environment and the role of the switch in the affected communication.
- Determine if there are violations on the associated class suspected, and iterate as necessary.

For example, the listed behavior has been detected:

- Devices cannot communicate to other devices outside of their network, but can communicate locally.
- Impact has been isolated to routed communication outside of the VLAN, and the switch acts as the default gateway.
- A check of the hosts indicates they cannot ping the gateway. After a check of their ARP table, the entry for the gateway remains as Incomplete.
- All other hosts that have the gateway resolve have no communication issues. A check of CoPP on the switch that acts as the gateway indicates there are violations on copp-system-p-class-normal.

<#root>

```
class-map copp-system-p-class-normal (match-any)
match access-group name copp-system-p-acl-mac-dot1x
match protocol arp
set cos 1
police cir 1400 kbps , bc 32000 bytes
module 1 :
transmitted 3292445628 bytes;

dropped 522023852 bytes;
```

- Additionally, multiple command checks show the drops are actively on the increase.
- These violations can cause legitimate ARP traffic to be dropped, which leads to a denial of services behavior.

It is important to highlight that CoPP isolates the impact to the traffic associated with the specific class, which in this example are ARP and copp-system-p-class-normal. Traffic related to other classes, such as OSPF, BGP is not be dropped by CoPP, as they fall within a different class entirely. If left unchecked, ARP issues can cascade into other problems, which can affect protocols that rely on it to begin with. For example, if an ARP cache times out and is not refreshed due to excessive violations, a TCP session such as BGP, can terminate.

- Control plane checks are advised to be performed, such as Ethanalyzer, CPU-mac in-band stats, and CPU process to isolate the matter further.

Ethanalyzer

As traffic policed by CoPP is associated only with CPU-bound traffic, one of the most important tools is the Ethanalyzer. This tool is a Nexus implementation of TShark and allows traffic sent and received by the supervisor to be captured and decoded. It can also use filters that are based on different criteria, such as protocols or header information, thus becomes an invaluable tool to determine traffic sent and received by the CPU.

The recommendation is to first examine the ARP traffic seen by the supervisor when the Ethanalyzer tool is run directly on the terminal session or sent to a file for analysis. Filters and limits can be defined to focus the capture into a specific pattern or behavior. To do this, add flexible display filters.

A common misconception is the Ethanalyzer captures all traffic that traverses through the switch. Data plane traffic, between hosts, is switched or routed by the hardware ASICs between data ports does not require CPU involvement and thus it is not normally seen by the Ethanalyzer capture. To capture data plane traffic, other tools, such as ELAM or SPAN are advised to be used. For example, to filter ARP, use the command:

```
ethanalyzer local interface inband display-filter arp limit-captured-frames 0 autostop duration 60 > arpcpu
```

Important configurable fields:

- interface inband - refers to traffic directed to the SUP
- display-filter arp - refers to the tshark filter applied, most Wireshark filters are accepted
- limit-captured-frames 0 - refers to the limit, 0 equates to unlimited, until stopped by another parameter or stopped manually by Ctrl+C
- autostop duration 60 - refers to the Ethanalyzer stop after 60 seconds, thus it creates a snapshot of 60 seconds of ARP traffic seen on the CPU

The Ethanalyzer output is redirected to a file on the bootflash with > arpcpu, to be manually processed. After 60 seconds, the capture completes, and the Ethanalyzer terminates dynamically, and the file arpcpu is on the bootflash of the switch, which can then be processed to extract the top talkers. For example:

```
show file bootflash:arpcpu | sort -k 3,5 | uniq -f 2 -c | sort -r -n | head lines 50

669 2022-05-10 10:29:50.901295 28:ac:9e:ad:5e:47 -> ff:ff:ff:ff:ff:ff ARP Who has 10.1.1.1? Tell 10.1.1
668 2022-05-10 10:29:50.901295 28:ac:9e:ad:5e:43 -> ff:ff:ff:ff:ff:ff ARP Who has 10.2.1.1? Tell 10.2.1
668 2022-05-10 10:29:50.901295 28:ac:9e:ad:5e:41 -> ff:ff:ff:ff:ff:ff ARP Who has 10.3.1.1? Tell 10.3.1
```

This filter is sorted based on: the source and destination columns, then the unique matches found (but ignores the date column), counts the instances and adds the number seen, and finally sorts top-to-bottom, based on count, and displays the first 50 results.

In this lab example, in 60 seconds, over 600 ARP packets were received from three devices, which have been identified as the suspected offender devices. The first column on the filter details the number of instances for this event were seen on the capture file in the specified duration.

It is important to understand the Ethanalyzer tool acts on the in-band driver,which is essentially the communication into the ASIC. In theory, the packet needs to pass through the kernel and the packet manager to be handed off to the associated process itself. CoPP and HWRL act before the traffic is seen on the Ethanalyzer. Even if violations are actively on the increase, some traffic still passes through and is conformed within the police rate, which helps provide insight into the traffic flows punted to the CPU. It is an important distinction, as traffic seen on the Ethanalyzer is NOT the traffic that violated the CIR and was dropped.

The Ethanalyzer can also be used in an open fashion, without any display filter or capture filter specified to catch all relevant SUP traffic. This can be used as an isolation measure as part of the approach to troubleshoot it.

For additional details and use of the Ethanalyzer, refer to the TechNote:

Ethanalyzer on Nexus 7000 Troubleshooting Guide

Using Ethanalyzer on Nexus platform for control-plane and data-plane traffic analysis

Note: Nexus 7000, prior to 8.X code release, can only perform Ethanalyzer captures through the admin VDC, which encompass SUP-bound traffic from all VDCs. VDC-specific Ethanalyzer is present in 8.X codes.

CPU-MAC In-band Stats

The in-band stats associated with CPU-bound traffic keep relevant statistics of in-band TX/RX CPU traffic. These statistics can be checked with the command: show hardware internal cpu-mac inband stats, which provides insight into the current rate and peak rate statistics.

```
show hardware internal cpu-mac inband stats`
================ Packet Statistics ======================
Packets received: 363598837
Bytes received: 74156192058
Packets sent: 389466025
Bytes sent: 42501379591
Rx packet rate (current/peak): 35095 / 47577 pps
Peak rx rate time: 2022-05-10 12:56:18
Tx packet rate (current/peak): 949 / 2106 pps
Peak tx rate time: 2022-05-10 12:57:00
```

As a best practice, it is advised that a baseline is created and tracked because due to the the role of the switch and the infrastructure, output of the **show hardware internal cpu-mac inband stats** varies significantly. In this lab environment, the usual values and historical peaks are typically no greater than a few hundred pps, and thus this is abnormal. The command **show hardware internal cpu-mac inband events** is also useful as a historical reference, because it contains data related to the peak use and the time it was detected.

Process CPU

The Nexus switches are Linux-based systems, and the Nexus Operating System (NXOS) takes advantage of CPU preemptive scheduler, multitasking, and multithreading of its respective cores architecture, to provide fair access to all processes, and thus spikes are not always indicative of a problem. However, if sustained traffic violations are seen, it is likely the associated process is also heavily used and appears as a top resource under the CPU outputs. Take multiple snapshots of the CPU processes to verify high use of a particular process by the use of: **show processes cpu sort | exclude 0.0 or show processes cpu sort | grep <process>.**

The process CPU, in-band stats, and Ethanalyzer verifications provide insight on the processes and traffic currently processed by the supervisor and help isolate on-going instability on control plane traffic that can cascade into data plane issues. It is important to understand that CoPP is a protection mechanism. It is reactionary because it only acts on traffic punted to the SUP. It is designed to safeguard the integrity of the supervisor by the discard of traffic rates, which exceed the expected ranges. Not all drops indicate a problem or require intervention, as their importance relates to the specific CoPP class and the verified impact, based on the infrastructure and network design. Drops due to sporadic burst events do not translate into impact, as protocols have build-in mechanisms, such as keepalive and retries that can deal with transient events. Keep the focus on sustained events or abnormal events beyond established baselines. Remember that CoPP must adhere to the protocols and features specific to the environment and must be monitored and continuously iterated upon to fine tune it, based on scalability needs as they evolve. If drops occur, determine if CoPP dropped traffic unintentionally or in response to a malfunction or attack. In either event, analyze the situation and evaluate the need to intervene by analysis of the impact and corrective measure on the environment, which can be outside of the scope of the switch itself.

Additional information

Recent platforms/codes, can have the ability to perform a SPAN-to-CPU, by the mirror of a port and punt of the data plane traffic to the CPU. This is normally heavily rate-limited by the hardware rate-limit and CoPP. Careful use of the SPAN to CPU is advised, and is outside the scope of this document.

Refer to the Tech Note listed for more information on this feature:

Nexus 9000 Cloud Scale ASIC NX-OS SPAN-to-CPU Procedure