Possibilities

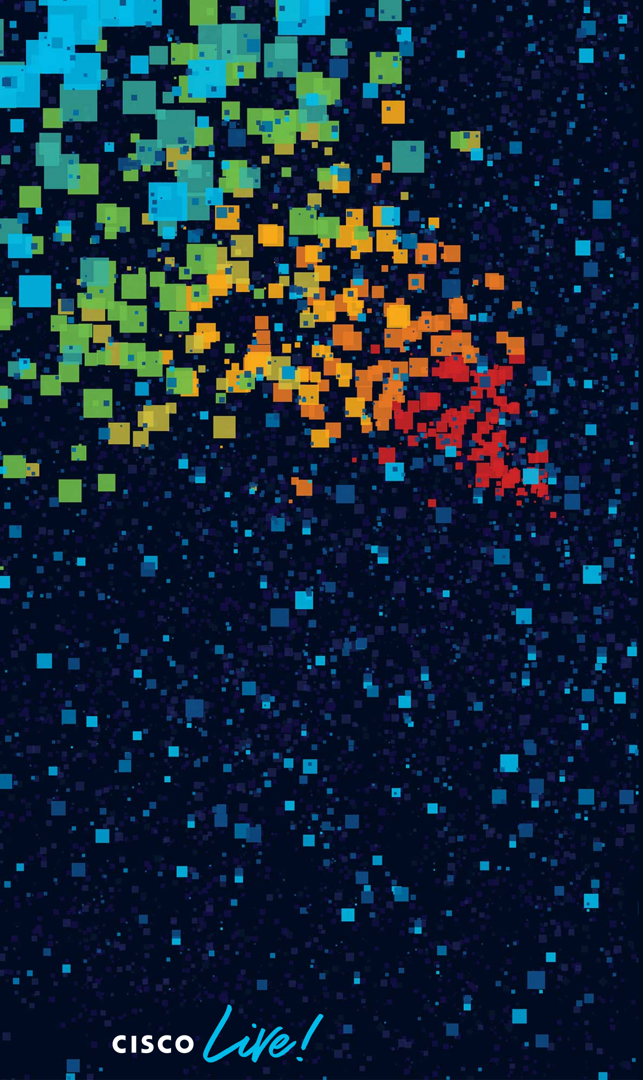
#CiscoLive

Troubleshoot packet loss on ASR9K

Scenarios and case studies

Sameer R. Avadhani
Technical Consulting Engineer, CCIE# 38105 – R&S, SP
@sameer1310

DGTL-TSCSPG-605



Agenda

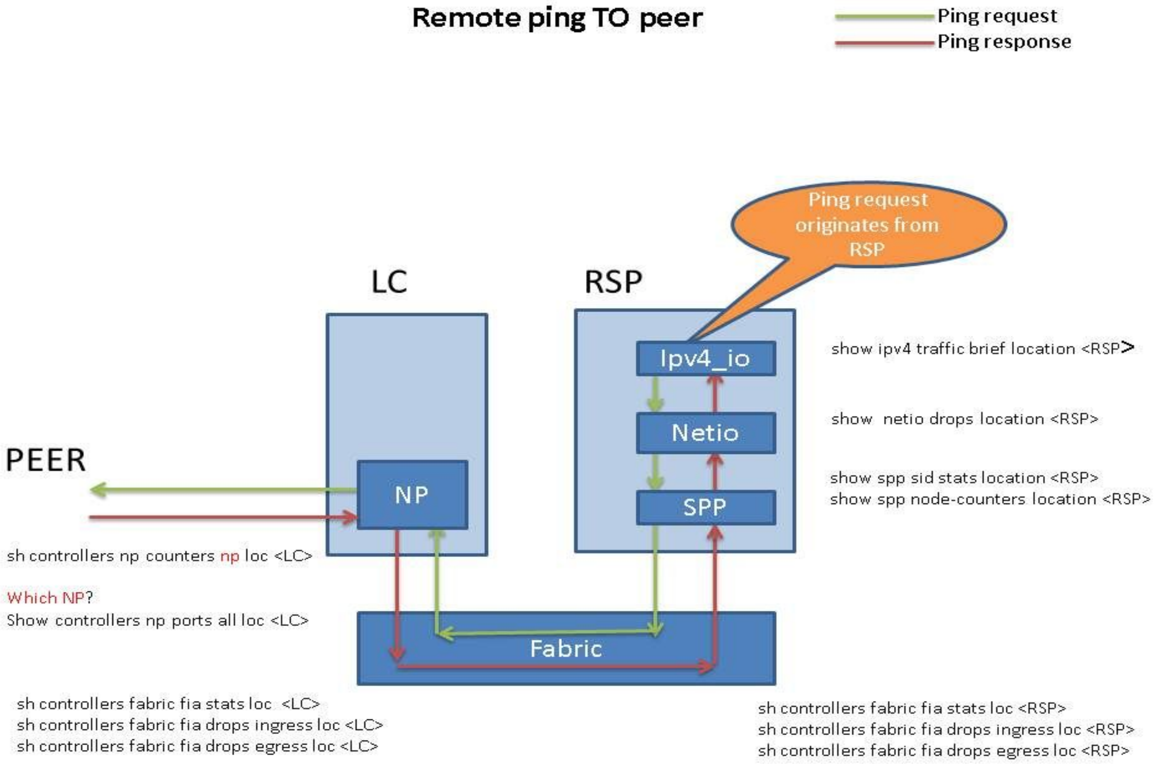
- Quick view of packet forwarding in ASR9K
- Troubleshooting steps for the most common type of packet loss issue
- Case studies
- Demo

Quick view of packet forwarding in ASR9K

Outgoing Ping Path

- Request: RSP CPU to outgoing interface/ controller
- Response: Incoming interface/ controller to RSP CPU

- **ICMP Request** - ipv4_io (RSP) -> Netio (RSP) -> SPP (RSP) -> Fabric -> NP -> outgoing i/f
- **ICMP Response** – incoming i/f -> NP -> Fabric -> SPP (RSP) -> Netio (RSP) -> ipv4_io (RSP)



show ipv4 traffic brief location <RSP>

show netio drops location <RSP>

show spp sid stats location <RSP>
show spp node-counters location <RSP>

sh controllers np counters np loc <LC>

Which NP?
Show controllers np ports all loc <LC>

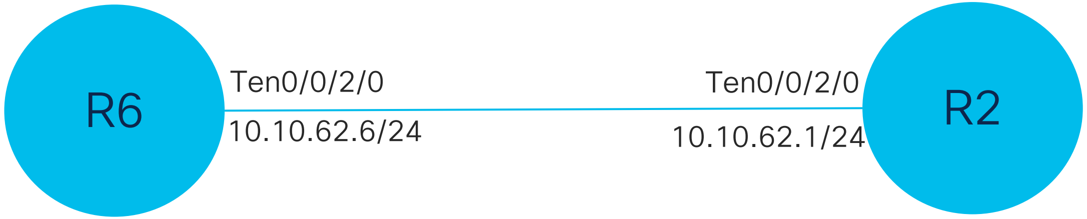
sh controllers fabric fia stats loc <LC>
sh controllers fabric fia drops ingress loc <LC>
sh controllers fabric fia drops egress loc <LC>

sh controllers fabric fia stats loc <RSP>
sh controllers fabric fia drops ingress loc <RSP>
sh controllers fabric fia drops egress loc <RSP>

Most common type of packet loss issue

Topology and scenario

- Below is a common case that gets opened for TAC to TS.
- Topology:



Scenario:

We ping from R2 to R6, and experience packet loss.

What are the steps to be performed to TS this issue?

What tools can we use to troubleshoot this issue?

Most common type of packet loss issue

Troubleshooting steps

Step 1) We open 2 sessions to the same router.

a) On one session, we initiate the ping to the peer.

b) On second session, we monitor the following output:

Before pinging

```
RP/0/RSP0/CPU0:R2#show ipv4 traffic br | in echo
Mon May 18 19:21:58.671 UTC
Sent >> 7 echo request, 109 echo reply
        109 echo request, 5 echo reply
RP/0/RSP0/CPU0:R2#
```

After pinging

```
RP/0/RSP0/CPU0:R2#show ipv4 traffic br | in echo
Mon May 18 19:22:07.024 UTC
        12 echo request, 109 echo reply
Recd >> 109 echo request, 10 echo reply
RP/0/RSP0/CPU0:R2#
```

From the above capture, you can see, we sent 5 echo requests to R6, and received 5 echo replies back from R6, hence we do NOT see a packet loss.

Let’s say, if we observe packet loss, that is, we send 5 echo requests, and got only 4 back, then there are 3 possibilities:

1) the echo requests actually did not leave the box, or,
2) they were received by the peer, but were dropped, or,
3) all echo replies were sent by the peer, and this box (R2) dropped the echo replies.

Most common type of packet loss issue

Troubleshooting steps

Step 2) Let’s assume there was a packet loss, and we want to confirm our 3 theories, we will check the controller of R2:

```
# show controllers tenGigE 0/0/2/1 stats
```

The above command will dump a huge output, now we are interested in finding out, whether we sent out all 5 echo requests or not.

For the same, check the output counters, and see which packet size has non-increasing value. For example, I have packet size between 512 and 1023 as non-increasing counter.
Once again, we will have 2 sessions open, we will ping from one, and monitor the below command, on the other.

On 1st session: `RP/0/RSP0/CPU0:R2#ping 10.10.62.6 size 600`

On 2nd session, we will see 5 echo request packets going out:

Before pinging

```
RP/0/RSP0/CPU0:R2#show controllers tenGigE
0/0/2/0 stats | in Output pkts 512-1023 bytes
Mon May 18 19:45:03.472 UTC
    Output pkts 512-1023 bytes   = 11
RP/0/RSP0/CPU0:R2#
```

After pinging

```
RP/0/RSP0/CPU0:R2#show controllers tenGigE
0/0/2/0 stats | in Output pkts 512-1023 bytes
Mon May 18 19:45:12.501 UTC
    Output pkts 512-1023 bytes   = 16
RP/0/RSP0/CPU0:R2#
```

Most common type of packet loss issue

Troubleshooting steps

Let’s say you have lot of background traffic, and cannot collect those controller stats, then we have an alternative.

Step 3) ACL, probably the best and the easiest option used by TAC engineers.

a) Create an ACL on either side of the link, as follows:

R6

```
The ingress ACL is checking for incoming echo-
requests from R2:

ipv4 access-list test-in
20 permit icmp host 10.10.62.1 host 10.10.62.6 echo
30 permit ipv4 any any  <<< to ensure transit traffic
isn’t dropped.
!

The Egress ACL is checking for echo replies sent to
R2:

ipv4 access-list test-out
10 permit icmp host 10.10.62.6 host 10.10.62.1 echo-
reply
30 permit ipv4 any any
!
```

R2

```
My ingress ACL is checking for incoming echo-replies
from R6:

ipv4 access-list test-in
20 permit icmp host 10.10.62.6 host 10.10.62.1 echo-
reply
30 permit ipv4 any any
!

The Egress ACL is checking for echo-requests sent to
R6:

ipv4 access-list test-out
10 permit icmp host 10.10.62.1 host 10.10.62.6 echo
30 permit ipv4 any any
!
```

Most common type of packet loss issue

Troubleshooting steps

Step 3b) Applying the ACL, the correct way:

```
RP/0/RSP0/CPU0:R2#conf t
Mon May 18 20:09:15.145 UTC
RP/0/RSP0/CPU0:R2(config)#int tenGigE 0/0/2/0
RP/0/RSP0/CPU0:R2(config-if)#ipv4 access-group test-in ingress hardware-count
interface-statistics
RP/0/RSP0/CPU0:R2(config-if)#ipv4 access-group test-out egress hardware-count
interface-statistics
```

Using the key-words, “hardware-count” followed by, “interface-statistics” is imperative.

Without those key-words, the counters while using the show ipv4 access-lists command will not show the packet count.

More detailed explanation:

https://www.cisco.com/c/en/us/td/docs/routers/crs/software/crs_r4-2/addr_serv/configuration/guide/b_ipaddr_cg42crs/b_ipaddr_cg42crs_chapter_01.html

Most common type of packet loss issue

Troubleshooting steps

Step 3c) We open 2 sessions to the same router, on one session, we initiate the ping to the peer, on second session, we monitor the following output:

Before pinging

```
RP/0/RSP0/CPU0:R2#show access-lists test-out hardware
egress location 0/0/CPU0
Mon May 18 20:10:10.169 UTC
ipv4 access-list test-out
 10 permit icmp host 10.10.62.1 host 10.10.62.6 echo
 30 permit ipv4 any any (527 hw matches)
RP/0/RSP0/CPU0:R2#
```

After pinging

```
RP/0/RSP0/CPU0:R2#show access-lists test-out hardware
egress location 0/0/CPU0
Mon May 18 20:12:24.688 UTC
ipv4 access-list test-out
 10 permit icmp host 10.10.62.1 host 10.10.62.6 echo
(5 hw matches)  < sent 5 echo requests.
 30 permit ipv4 any any (3305 hw matches)
```

Before pinging

```
RP/0/RSP0/CPU0:R2#show access-lists test-in hardware
ingress location 0/0/CPU0
Mon May 18 20:19:27.700 UTC
ipv4 access-list test-in
 20 permit icmp host 10.10.62.6 host 10.10.62.1 echo-
reply
 30 permit ipv4 any any (12140 hw matches)
RP/0/RSP0/CPU0:R2#
```

After pinging

```
RP/0/RSP0/CPU0:R2#show access-lists test-in hardware
ingress location 0/0/CPU0
Mon May 18 20:19:27.700 UTC
ipv4 access-list test-in
 20 permit icmp host 10.10.62.6 host 10.10.62.1 echo-
reply (5 hw matches) < recd 5 echo replies.
 30 permit ipv4 any any (12140 hw matches)
RP/0/RSP0/CPU0:R2#
```

Most common type of packet loss issue

Troubleshooting steps

Step 3d) We open 2 sessions to the peer router, on one session, we initiate the ping to the peer, on second session, we monitor the following output:

Before pinging

```
RP/0/RSP0/CPU0:R6#show access-lists test-in hardware
ingress location 0/0/CPU0
Mon May 18 20:17:55.370 UTC
ipv4 access-list test-in
 20 permit icmp host 10.10.62.1 host 10.10.62.6 echo
 30 permit ipv4 any any (454 hw matches)
RP/0/RSP0/CPU0:R6#
```

```
RP/0/RSP0/CPU0:R6#show access-lists test-out
hardware egress location 0/0/CPU0
Mon May 18 20:36:41.083 UTC
ipv4 access-list test-out
 10 permit icmp host 10.10.62.6 host 10.10.62.1
echo-reply
 30 permit ipv4 any any (23645 hw matches)
RP/0/RSP0/CPU0:R6#
```

After pinging

```
RP/0/RSP0/CPU0:R6#show access-lists test-in hardware
ingress location 0/0/CPU0
Mon May 18 20:18:11.902 UTC
ipv4 access-list test-in
 20 permit icmp host 10.10.62.1 host 10.10.62.6 echo
(5 hw matches)  <<< echo requests from R2
 30 permit ipv4 any any (796 hw matches)
RP/0/RSP0/CPU0:R6#
```

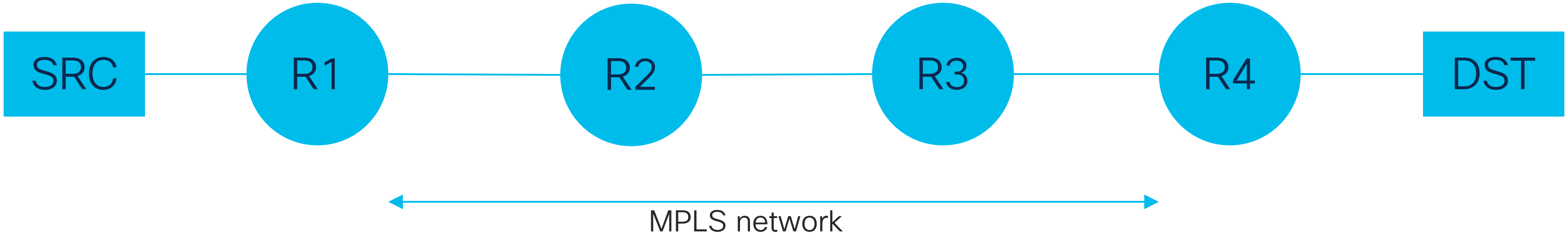
```
RP/0/RSP0/CPU0:R6#show access-lists test-out
hardware egress location 0/0/CPU0
Mon May 18 20:37:25.680 UTC
ipv4 access-list test-out
 10 permit icmp host 10.10.62.6 host 10.10.62.1
echo-reply (5 hw matches) <<< echo replies to R2
 30 permit ipv4 any any (24565 hw matches)
RP/0/RSP0/CPU0:R6#
```

Most common type of packet loss issue

Topology and scenario

- This is how we determine, which direction the packet loss is, and how to isolate the device with an issue.
- **Caveats to the discussion thus far:**
- All the above steps are applicable if we have a pain vanilla IP ckt/ network. **What if we have a MPLS based network, where packets are MPLS encapsulated, end to end?**

Example: We have the below topology:



Most common type of packet loss issue

Topology and scenario

- Here in this case, we cannot use ACL on an MPLS enabled link, as the packets from the source to destination will be MPLS encapsulated.
- We will have to ping hop by hop starting at the source to the destination.
- TAC will typically make use of NetFlow and check if the packets are forwarded or being dropped for a packet with source and destination in question.
- Please note, a user will have to use bulk ping of say 100,000 packets with timeout 0, for the netflow to catch these packets, due to the sampling limitations on netflow.
- Latest netflow config guide:

https://www.cisco.com/c/en/us/td/docs/routers/asr9000/software/asr9k-r7-1/netflow/configuration/guide/b-netflow-cg-asr9k-71x/b-netflow-cg-asr9k-70x_chapter_010.html

Check the appendix for a sample config and output of the verification of netflow.

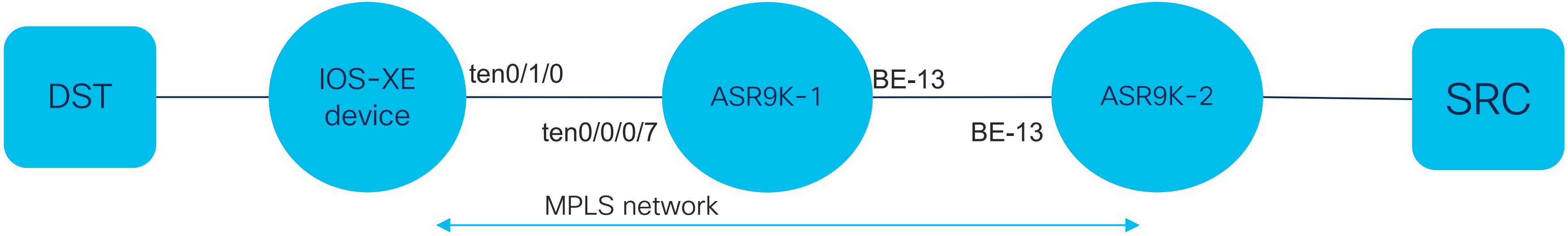
Case study

Case studies.

Case study 1

- Customer opened a TAC case complaining of a packet loss from SRC to DST, where some applications would be slow/ fail to respond during high peak hours.

Topology:



Scenario:

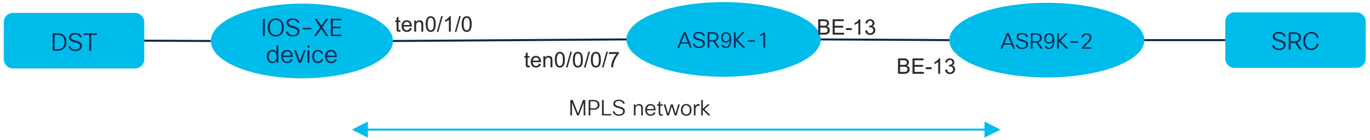
In many service providers/ enterprise networks, the design/ implementation plays a big part. This case study is a classic example of how a design flaw causes packet loss.

If you observe closely, the incoming traffic/ link capacity at ASR9K-1 is on a BE (multiple 10gig links), and outgoing to IOS-XE device is a single 10Gigs link.

How does this cause packet loss? What steps to take in order to identify the same?

Case studies.

Case study 1



Troubleshooting steps:

1) We started checking hop-by-hop, since we are focusing on ASR9K, we first look at ingress device that is ASR9K-2. We did a ping with a high count, end to end, and obviously, few packets dropped.

We check for:

a) Input drops on the incoming interface – found none.

b) Input errors/ CRC errors on the incoming interface – found none.

c) Any HW/ SW alarms on ingress/ egress LC – found none.

d) Output drops on the outgoing interface – found none.

e) Output errors/ CRC errors on the outgoing interface – found none.

f) Any NP drops/ errors on ingress and egress NP – found none.

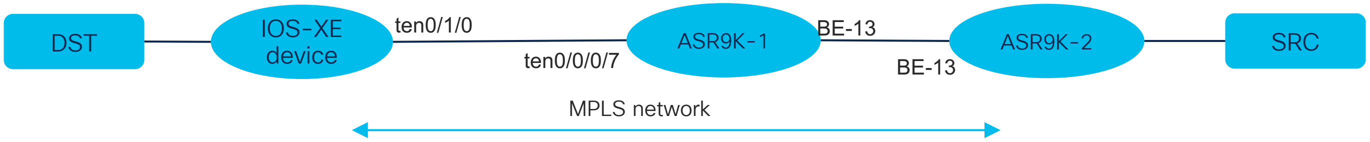
g) Any fabric drops/ errors on ingress and egress NP – found none.

ASR9K-2

2) Now, we cannot enable ACL, as this is a MPLS network, and we cannot configure Netflow, as this was a production environment.

Case studies.

Case study 1



Troubleshooting steps:

3) Now, since we are limited on resources, we will have to do the same checks on the next hop, that is ASR9K-1.

We started checking our list and when we come to point # g (from the previous slide), we see:

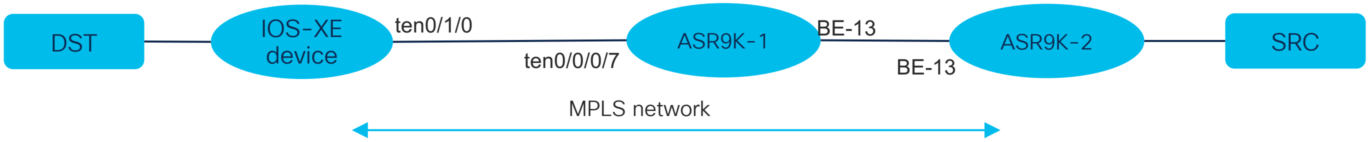
```
RP/0/RSP0/CPU0:ASR9K-1#sh controllers fabric fia drops ingress location 0/0/CPU0 | ex "
0"
Tue Dec  3 00:56:16.424

********** FIA-1 **********
Category: in_drop-0
                                  Hard Drop-2                    10400521
                                  WRED Drop-0                         814
                                  WRED Drop-2                   667376332   <<< These drops were
increasing constantly.

RP/0/RSP0/CPU0:ASR9K-1#
```

ASR9K-1

Case studies.

Case study 1

Troubleshooting steps:



WRED-Drops imply congestion, but where and how?

4) Now, let’s check the NP and FIA port mapping with the interfaces.

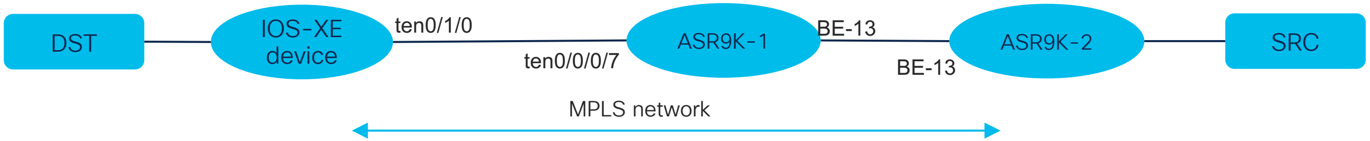
ASR9K-1

```
RP/0/RSP0/CPU0:ASR9K-1#show controllers np ports all location 0/0/cpu0
Sat Dec  7 01:12:17.724

                Node: 0/0/CPU0:
----------------------------------------------------------------

NP Bridge Fia                       Ports
-- ------ --- ---------------------------------------------------
0  0      0   TenGigE0/0/0/5      <<<< Ingress interface, part of BE-13
1  0      0   TenGigE0/0/0/3
2  1      0   TenGigE0/0/0/4      <<<< Ingress interface, part of BE-13
3  1      0   TenGigE0/0/0/2
4  2      1   TenGigE0/0/0/0
5  2      1   TenGigE0/0/0/1
6  3      1   TenGigE0/0/0/7      <<<< Egress interface to DST, NP is 6, FIA is FIA-1.
7  3      1   TenGigE0/0/0/6
```

Now, we look at BE-13, the incoming interface config. We see that, interfaces ten0/0/0/4, ten0/0/0/5, ten0/1/0/4 and ten0/1/0/5 are the BE members.

Case studies.

Case study 1



Troubleshooting steps:

5) Now, we check the q-depth, to find out where do we have the congestion.

```
RP/0/RSP0/CPU0:ASR9K-1#show controllers fabric fia q-depth location 0/0/cpu0
<snip>
 ********** FIA-1 **********
Category: q_stats_a-1
Voq         ddr             pri             pktcnt
19          0               2               1024                 >> Packets congested at VoQ 19

RP/0/RSP0/CPU0:ASR9K-1#
```

Let's convert this 19 in a hex value. That is 0x13, we will need this for further verification.

6) Now, we need to verify what interface the VoQ is mapped to.

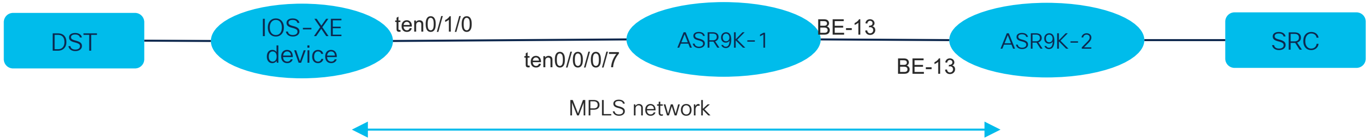
```
RP/0/RSP0/CPU0:ASR9K-1#show controllers pm location 0/0/cpu0 | in "switch_fabric|Ifname"
Sat Dec  7 01:24:13.469
Ifname(2): TenGigE0_0_0_7, ifh: 0x2000140 :
switch_fabric_port 0x13                                          >> 19 on VOQ o/p (hex value of 0x13)

RP/0/RSP0/CPU0:ASR9K-1#
```

We see that, the backpressure is seen on ten0/0/0/7. If you recall, this interface is mapped to NP6.

Case studies.

Case study 1

Troubleshooting steps:

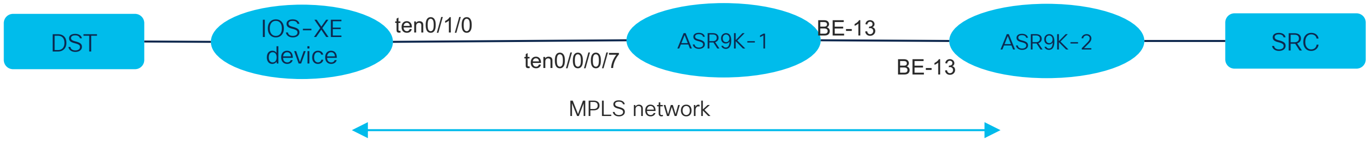


- Now, upon understanding the customer design further, and analyzing traffic flows, we found out that, 6 interfaces on the LC 0/0/cpu0, are ingress, and only 1 interface is egress for the same DST.
- These 6 interfaces are in addition to the interfaces from different LCs, remember, our BE-13 also has interfaces from 0/1/cpu0.

Hence, there was tremendous amount of congestion on this ASR9K device, for this traffic flow. This phenomenon is known as backpressure, where the egress NP, that is NP6 in this case, signals other NPs to delay sending traffic to it.

Case studies.

Case study 1



Troubleshooting steps:

In newer SW versions and Typhoon and beyond family of LCs, you will get a better view like below:

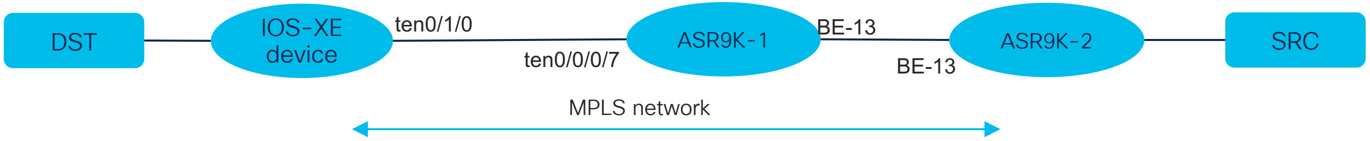
```
RP/0/RSP0/CPU0:R8#show controllers fabric fia q-depth location 0/0/cpu0
Mon May 25 16:57:38.836 UTC

 ********** FIA-0 **********
Category: q_stats_a-0
Voq         ddr              pri              pktcnt     Slot_FIA_NP
103         0                2                1394       LC0_1_1

 ********** FIA-0 **********
```

We at least know the NP and FIA, now to obtain the interface mapped to the NP, we will use the below command:

```
RP/0/RSP0/CPU0:R8#sho controllers pm vqi loc 0/0/cpu0
Tue May 26 03:26:17.981 UTC
Platform-manager VQI Assignment Information
      Interface Name        | ifh Value | VQI | NP#
--------------------------------------------------
           <snip>
           TenGigE0_0_2_3 | 0x4000280 | 103 |    1   <<<< 103 is the VoQ, that was congested.
RP/0/RSP0/CPU0:R8#
```

Case studies.

Case study 1



Solution:

We suggested the customer to add more interfaces either from different LCs or the same LC, to the egress interface ten0/0/0/7 and bundle them together to ensure there is enough BW for all the incoming traffic/ requests from the source.

References:

In depth understanding of backpressure:

https://www.ciscolive.com/c/dam/r/ciscolive/us/docs/2014/pdf/BRKSPG%202904.pdf

https://community.cisco.com/t5/service-providers-documents/asr9000-xr-understanding-qos-default-marking-behavior-and/ta-p/3128709

https://community.cisco.com/t5/xr-os-and-platforms/asr9922-qos-vqi-voq/td-p/2616291

In depth details on packet drops and understanding NP counters:

https://community.cisco.com/t5/service-providers-documents/asr9000-xr-troubleshooting-packet-drops-and-understanding-np/ta-p/3126715

Case studies.

Other examples or case studies

Typical issues seen by TAC are due to the following:

1) Layer-1 issues, like bad optic/ cables/ the device's HW/ controller itself.

2) Introduction of an intermediate device such as MUXs or any layer-1 pass through equipment, which are typically 3rd party vendors in LEC or ISP.

3) Cases where customer had ACLs blocking IPv6 multicast traffic, and experienced loss of hello packets for IGP neighborship.

4) Mismatching MTUs, somewhere in the network, also lead to keepalive packets to be lost for certain protocols and can cause control plane failure between routers.

Demo

Thank you

cisco

Possibilities

#CiscoLive