

Troubleshooting Input Discards on M3 Series Linecards

Introduction

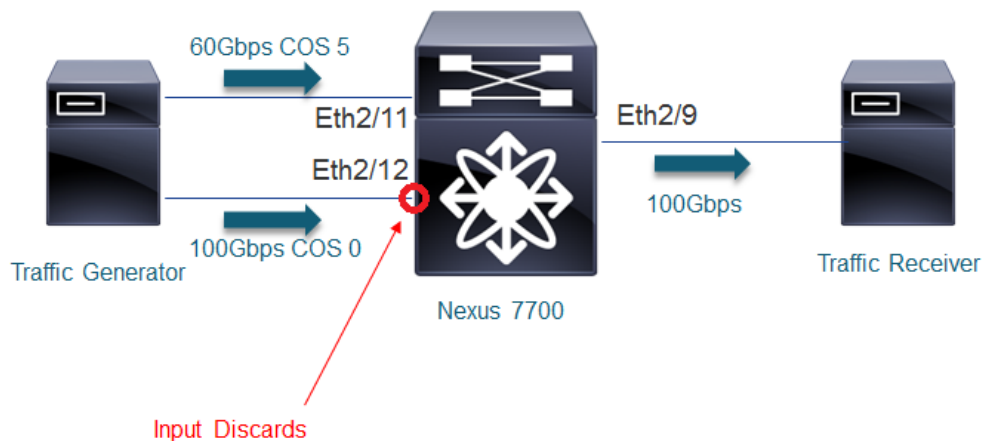
This document explains the procedure to troubleshoot input discard on M3 Line card interfaces.

Contributed by Manoj Shukla, Nikolay Kartashev, Ray Kraft, Cisco TAC Engineers.

Topology

Lab setup consists of Nexus 7700 with M3-series Line card. The following two flows are sent through the network.

Ingress Port	Egress Port	Traffic Rate	COS Marking
Eth2/11	Eth2/9	60Gbps	5
Eth2/12	Eth2/9	100Gbps	0



The following output shows hardware configuration of Nexus 7710 switch.

```
Angeles2# show module
Mod  Ports  Module-Type                Model                Status
----  -
2    12     40/100 Gbps Ethernet Module  N77-M312CQ-26L     ok
```

```
9    0    Supervisor Module-2    N77-SUP2E    active *
10   0    Supervisor Module-2    N77-SUP2E    ha-standby
```

```
Mod  Sw          Hw
---  -
2    8.0(1)       0.8
9    8.0(1)       0.308
10   8.0(1)       0.308
```

Interfaces used for testing are 100Gbps, and have a basic layer 2 configuration.

```
Angeles2# show running-config interface Ethernet2/9, Ethernet2/11, Ethernet2/12
```

```
version 8.0(1)

interface Ethernet2/9
  description --Egress-Interface--
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 100
  no shutdown

interface Ethernet2/11
  description --Ingress-Interface--
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 100
  no shutdown

interface Ethernet2/12
  description --Ingress-Interface--
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 100
  no shutdown
```

```
Angeles2#
```

Problem Description

Input discards keep incrementing on M3 Line card-based interface Ethernet 2/12.

```
Angeles2# show interface ethernet 2/12
Ethernet2/12 is up
admin state is up, Dedicated Interface
Hardware: 40000/100000 Ethernet, address: 00b0.e134.74e0 (bia 00b0.e134.74e0)
MTU 9216 bytes, BW 100000000 Kbit, DLY 10 usec
reliability 255/255, txload 1/255, rxload 251/255
Encapsulation ARPA, medium is broadcast
Port mode is trunk
full-duplex, 100 Gb/s, media type is 100G
Beacon is turned off
Auto-Negotiation is turned on
```

```

Input flow-control is off, output flow-control is off
Auto-mdix is turned off
Rate mode is dedicated
Switchport monitor is off
EtherType is 0x8100
EEE (efficient-ethernet) : n/a
admin fec state is auto, oper fec state is cl91
Last link flapped 3d15h
Last clearing of "show interface" counters never
7 interface resets
Load-Interval #1: 30 seconds
30 seconds input rate 98686031736 bits/sec, 8231697 packets/sec
30 seconds output rate 848 bits/sec, 0 packets/sec
input rate 98.69 Gbps, 8.23 Mpps; output rate 848 bps, 0 pps
Load-Interval #2: 5 minute (300 seconds)
300 seconds input rate 98684727088 bits/sec, 8230526 packets/sec
300 seconds output rate 352 bits/sec, 0 packets/sec
input rate 98.68 Gbps, 8.23 Mpps; output rate 352 bps, 0 pps
RX
63479707145 unicast packets 0 multicast packets 0 broadcast packets
63479738967 input packets 95219893024500 bytes
0 jumbo packets 0 storm suppression packets
0 runts 0 giants 0 CRC/FCS 0 no buffer
0 input error 0 short frame 0 overrun 0 underrun 0 ignored
0 watchdog 0 bad etype drop 0 bad proto drop 0 if down drop
0 input with dribble 38058645989 input discard
0 Rx pause
TX
0 unicast packets 5633 multicast packets 0 broadcast packets
5629 output packets 866261 bytes
0 jumbo packets
0 output error 0 collision 0 deferred 0 late collision
0 lost carrier 0 no carrier 0 babble 0 output discard
0 Tx pause

```

Angeles2#

Troubleshooting Steps

The following is the command to find congested Egress interface, which causes input discards increment on a particular Ingress interface.

```
show hardware queuing drops [module <module #>]
```

However, before using this command, conversion of ingress front panel port must be performed according to the following table

Front Panel Port	VQ Drops Source Interface
1	1-4
2	5-8

3	8-12
4	13-16
5	17-20
6	21-24
7	25-28
8	29-32
9	33-36
10	37-40
11	41-44
12	45-48

Here, for Ingress interface Ethernet 2/12 where input discards are registered, VQ Drops Source Interface range would be Ethernet 2/45-48.

In the following output, VQ Drops Source Interface range Ethernet 2/37-40,45-48 corresponds to front panel ports Ethernet 2/10 and Ethernet 2/12 (they reside on the same ASIC, hence both show up in the output), and Egress Port causing congestion is Ethernet 2/9.

```
Angeles2# show hardware queuing drops
```

```
-----|
| Executed at : 2017-01-24 23:23:04.784 |
|-----|
```

VQ Drops

```
-----|
| Source | Egresss | Congested | Dropped |
| Interface | Port | Queue | Packet |
|-----|
| Eth2/37-40,45-48 | Eth2/9 | 7 | 0000040507743599 |
|-----|
```

Egress Buffer Drops

```
-----|
| Output | EB Drop pkts |
| Interface |
|-----|
```

Span Drops

```
-----|
```

input Interface	SPAN session	span drop pkts

Tip: you can clear hardware counters in the above output with **clear statistics module-all device qengine** command

You can narrow down which COS value(s) are dropped as ingress discards on the interface. For this you can find which queuing class drops are occurring

```
Angeles2# show policy-map interface ethernet 2/12 input
```

```
Global statistics status : enabled
```

```
Ethernet2/12
```

```
Service-policy (queuing) input: default-8e-4q8q-in-policy
SNMP Policy Index: 301997241
```

```
Class-map (queuing): 8e-4q8q-in-q1 (match-any)
  queue-limit percent 10
  bandwidth percent 49
  queue dropped pkts : 0
  queue dropped bytes : 0
  queue transmit pkts: 0 queue transmit bytes: 0
```

```
Class-map (queuing): 8e-4q8q-in-q-default (match-any)
  queue-limit percent 88
  bandwidth percent 49
  queue dropped pkts : 50274556148
  queue dropped bytes : 0 current depth bytes: 319794136
  queue transmit pkts: 83820000221 queue transmit bytes: 125730001497404
```

```
Class-map (queuing): 8e-4q8q-in-q3 (match-any)
  queue-limit percent 1
  bandwidth percent 1
  queue dropped pkts : 0
  queue dropped bytes : 0
  queue transmit pkts: 0 queue transmit bytes: 0
```

```
Class-map (queuing): 8e-4q8q-in-q4 (match-any)
  queue-limit percent 1
  bandwidth percent 1
  queue dropped pkts : 0
  queue dropped bytes : 0
  queue transmit pkts: 0 queue transmit bytes: 0
```

```
Angeles2#
```

Since drops are registered in class-map 8e-4q8q-in-q-default, you can check which COS values are mapped to this class-map on ingress

```
Angeles2# show queuing interface ethernet 2/12 module 2
```

```
Egress Queuing for Ethernet2/12 [System]
```

```
-----
Template: 8Q8E
```

```
-----
```

Queue	Group	Bandwidth%	PrioLevel	Shape%	CoSMap
8e-4q8q-out-q4	3	14	-	-	4
8e-4q8q-out-q5	4	14	-	-	3
8e-4q8q-out-q6	5	14	-	-	2
8e-4q8q-out-q7	6	14	-	-	1
8e-4q8q-out-q1	0	-	High	-	5
8e-4q8q-out-q2	1	14	-	-	7
8e-4q8q-out-q3	2	14	-	-	6
8e-4q8q-out-q-default	7	14	-	-	0

Ingress Queuing for Ethernet2/12 [System]

```
-----
Trust: Trusted
DSCP to Ingress Queue : Disabled
```

```
-----
```

Queue	Group	Qlimit%	IVL	CoSMap	EXPMMap
8e-4q8q-in-q-default	1	88	0	0-4,6	0-4,6
8e-4q8q-in-q1	0	10	2	5	5
8e-4q8q-in-q4	3	1	4	7	7
8e-4q8q-in-q3	2	1	-	-	-

Angeles2#

Note that no input discards are registered on interface Ethernet2/11, as it sends CoS 5 priority traffic.

```
Angeles2# show interface ethernet 2/11
Ethernet2/11 is up
admin state is up, Dedicated Interface
  Hardware: 40000/100000 Ethernet, address: 00b0.e134.74dc (bia 00b0.e134.74dc)
  MTU 1500 bytes, BW 100000000 Kbit, DLY 10 usec
  reliability 255/255, txload 1/255, rxload 150/255
  Encapsulation ARPA, medium is broadcast
  Port mode is trunk
  full-duplex, 100 Gb/s, media type is 100G
  Beacon is turned off
  Auto-Negotiation is turned on
  Input flow-control is off, output flow-control is off
  Auto-mdix is turned off
  Rate mode is dedicated
  Switchport monitor is off
  EtherType is 0x8100
  EEE (efficient-ethernet) : n/a
    admin fec state is auto, oper fec state is cl91
  Last link flapped 05:49:41
  Last clearing of "show interface" counters never
  7 interface resets
  Load-Interval #1: 30 seconds
    30 seconds input rate 59198977752 bits/sec, 4938868 packets/sec
    30 seconds output rate 896 bits/sec, 0 packets/sec
    input rate 59.20 Gbps, 4.94 Mpps; output rate 896 bps, 0 pps
  Load-Interval #2: 5 minute (300 seconds)
    300 seconds input rate 59211977616 bits/sec, 4938427 packets/sec
    300 seconds output rate 528 bits/sec, 0 packets/sec
    input rate 59.21 Gbps, 4.94 Mpps; output rate 528 bps, 0 pps
  RX
    61330589321 unicast packets  0 multicast packets  0 broadcast packets
```

```
61330601836 input packets 91996083141000 bytes
0 jumbo packets 0 storm suppression packets
0 runts 0 giants 0 CRC/FCS 0 no buffer
0 input error 0 short frame 0 overrun 0 underrun 0 ignored
0 watchdog 0 bad etype drop 0 bad proto drop 0 if down drop
0 input with dribble 0 input discard
0 Rx pause
TX
0 unicast packets 9005 multicast packets 0 broadcast packets
9004 output packets 1385289 bytes
0 jumbo packets
0 output error 0 collision 0 deferred 0 late collision
0 lost carrier 0 no carrier 0 babble 0 output discard
0 Tx pause
```

Angeles2#

You can confirm receipt of CoS 5 priority traffic on interface Ethernet 2/11

Angeles2# **show policy-map interface Ethernet 2/11 input**

Global statistics status : enabled

Ethernet2/11

```
Service-policy (queuing) input: default-8e-4q8q-in-policy
SNMP Policy Index: 301995853
```

```
Class-map (queuing): 8e-4q8q-in-q1 (match-any)
  queue-limit percent 10
  bandwidth percent 49
  queue dropped pkts : 0
  queue dropped bytes : 0 current depth bytes: 9096
queue transmit pkts: 65991127103 queue transmit bytes: 98986691355220
```

```
Class-map (queuing): 8e-4q8q-in-q-default (match-any)
  queue-limit percent 88
  bandwidth percent 49
  queue dropped pkts : 0
  queue dropped bytes : 0
  queue transmit pkts: 0 queue transmit bytes: 0
```

```
Class-map (queuing): 8e-4q8q-in-q3 (match-any)
  queue-limit percent 1
  bandwidth percent 1
  queue dropped pkts : 0
  queue dropped bytes : 0
  queue transmit pkts: 0 queue transmit bytes: 0
```

```
Class-map (queuing): 8e-4q8q-in-q4 (match-any)
  queue-limit percent 1
  bandwidth percent 1
  queue dropped pkts : 0
  queue dropped bytes : 0
  queue transmit pkts: 0 queue transmit bytes: 0
```

Angeles2#

With the following output you can confirm mapping of CoS 5 to Ingress Queue 8e-4q8q-in-q1.

```
Angeles2# show queuing interface ethernet 2/11 module 2
```

```
Egress Queuing for Ethernet2/11 [System]
```

```
-----  
Template: 8Q8E  
-----
```

Queue	Group	Bandwidth%	PrioLevel	Shape%	CoSMap
8e-4q8q-out-q4	3	14	-	-	4
8e-4q8q-out-q5	4	14	-	-	3
8e-4q8q-out-q6	5	14	-	-	2
8e-4q8q-out-q7	6	14	-	-	1
8e-4q8q-out-q1	0	-	High	-	5
8e-4q8q-out-q2	1	14	-	-	7
8e-4q8q-out-q3	2	14	-	-	6
8e-4q8q-out-q-default	7	14	-	-	0

```
Ingress Queuing for Ethernet2/11 [System]
```

```
-----  
Trust: Trusted
```

```
DSCP to Ingress Queue : Disabled  
-----
```

Queue	Group	Qlimit%	IVL	CoSMap	EXPMap
8e-4q8q-in-q-default	1	88	0	0-4,6	0-4,6
8e-4q8q-in-q1	0	10	2	5	5
8e-4q8q-in-q4	3	1	4	7	7
8e-4q8q-in-q3	2	1	-	-	-

```
Angeles2#
```

Determining Traffic Sources of an Egress Port

Note: Due to an architecture design limitation this can only accurately indentify ingress SOCs and cannot provide ingress port granularity.

With the M3 series line card in 8.x we have additional CLIs that can be leveraged to determine all sources of traffic for a given egress SOC in addition to LDI/LQI mappings. During this test Eth2/11 was shutdown, so the only flow was Eth2/12 >> Eth2/9.

The command below shows all of the VQIs that the egress port is receiving traffic from. Under 'SRC PORT' the first number is a VQI in hex and the second number after '/' is the same VQI in decimal. Per the legend at the bottom of the output there are 3 possible bits that can be set; 'A' for 'Arrival' indicates that this SOC is receiving traffic from that VQI. Additionally, you must increment the VQI per the columns 0-F at the top of the table. If 'A' was set in column 0 of the first entry this would correspond to VQI 0x90/144. In this case the VQIs of interest are 0x93/147 and 0x101/257.


```
module-2# show hardware internal qengine inst 5 table vrf_rfi_sps
```

```
-----
| VRF_RFI_SPS (physical layout) for SLF VQ
| Inst 5; port(s) 32-35,40-43
|
-----
SRC PORT    0    1    2    3    4    5    6    7    8    9    A    B    C    D    E    F
-----
090/0144    --- --- --- -A --- --- --- --- --- --- --- --- --- --- --- ---
100/0256    --- --A --- --- --- --- --- --- --- --- --- --- --- --- --- ---
-----
C - Chico error or FFCFF ECC error; R - CRC error; A - Arrival
```

Next check the vqi-map to map the VQI to the SLOT, LQI, NUM (ASIC instance). SLOT is 0 based so this must be incremented by one. This shows that VQI 147 maps to a port on module/slot 2, on ASIC instance 4, with LQI 12. VQI 257 maps to slot 9 which is the SUP in this chassis.

```
module-2# show hardware internal qengine vqi-map | i i vqi|147
VQI  SLOT LDI LQI DFI FPOE NUM XMSK ASC ASC NO VQI HDR CHNK VQI PRT FEATURE
147  1    9   12 10  40   4   3ff SLF 2  no  DAT V5  0   100 100 880084 <<<<<
Port on module 2
```

```
module-2# show hardware internal qengine vqi-map | i i vqi|257
VQI  SLOT LDI LQI DFI FPOE NUM XMSK ASC ASC NO VQI HDR CHNK VQI PRT FEATURE
257  8    1    1  192 256  1   3ff CLP 0  no  INB V4  0   10G 10G 14   <<<<<
SUP in slot 9
```

The front-panel-port-map can be used to map the LQI to the specific port however due to a design limitation with this architecture the result is inconsistent. Using the NUM (ASIC INST) value and the LQI (ASIC PORT) the 147 VQI entry should map to Eth2/10.

```
module-2# sh hard internal qengine front-panel-port-map
FP PORT    ASIC INST    ASIC PORT
(1-based)  (0-based)    (0-based)
-----
1           3             0
2           0             12
3           3             12
4           0             0
5           4             0
6           1             12
7           4             12
8           1             0
9           5             0
10          2             12 <<<<<<<<<< SLOT 2, ASIC INST 4, LQI 12 maps to FP (front
panel port) 10
11          5             12
12          2             0  <<<<<<<<<< Eth2/12 shares the same ingress ASIC instance 4
```

EthPM can be used as a cross reference of the VQI to verify findings. In this case we can see that Eth2/10 does not correspond to VQI 147. “Sport Port” used by SPS comes from the DC3 header and is used for unicast reordering; it is assigned per ASIC, not per FP port. All traffic originated from the same ingress ASIC carries the same SOURCE PORT value, which is the first VQI

assigned to the ASIC. So these commands can be used to accurately identify the ingress ASIC but not necessarily the correct ingress port. In this case there are only 2 ports per SOC/ASIC (Eth2/10 and Eth2/12), so by process of elimination we know that the actual sender is Eth2/12.

```
Angeles2# show system internal ethpm info interface e2/10 | i i vqi  
LTL(0xb5e), VQI(0x93), LDI(0x9), IOD(0x193)
```

```
Angeles2# dec 0x93  
147
```

```
Angeles2# show int e2/10 | head line 1  
Ethernet2/10 is down (XCVR not inserted)
```

```
Angeles2# show int e2/12 | he 1 1  
Ethernet2/12 is up
```