

瞭解Nexus 5600/6000的輸入丟棄並對其進行故障排除

目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[背景資訊](#)

[單播流量和緩衝](#)

[組播流量和緩衝](#)

[什麼原因導致輸入丟棄？](#)

[疑難排解案例](#)

[場景1.輸入丟棄](#)

[步驟1.使用輸入丟棄識別埠](#)

[步驟2. ASIC標識](#)

[步驟3.確定出口擁塞埠](#)

[場景2.使用HOLB輸入丟棄](#)

[HOLB緩解：啟用VOQ限制](#)

[HOLB緩解：流量分類](#)

[相關資訊](#)

簡介

本文說明如何對Cisco Nexus 5600/6000系列交換機上的輸入丟棄進行故障排除。

必要條件

需求

思科建議您瞭解Cisco Nexus 6000系列配置的基本知識。

採用元件

本文中的資訊係根據以下軟體和硬體版本：

- Cisco Nexus 6001
- 7.1(3)N1(1)

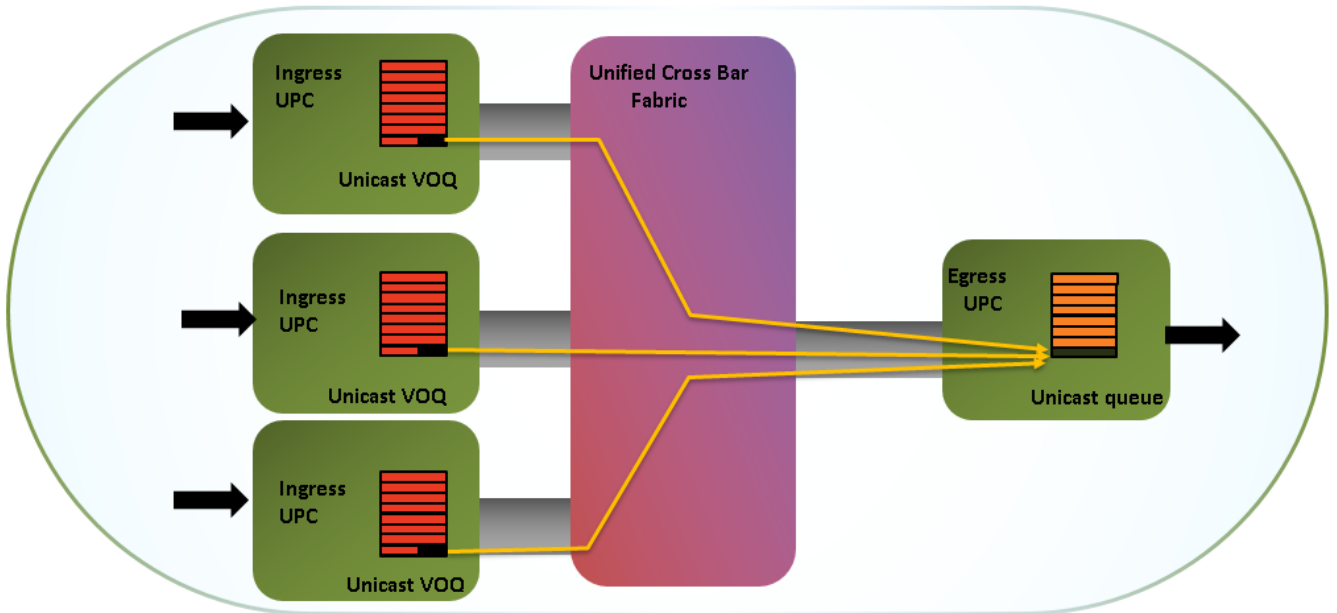
本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

背景資訊

輸入丟棄是超訂用出口埠的指示。這還意味著您很可能會在該特定連線埠上捨棄單點傳播流量。本文幫助您瞭解單播和組播流量在此平台上的緩衝方式，以及輸入丟棄是如何發生的以及緩解步驟。

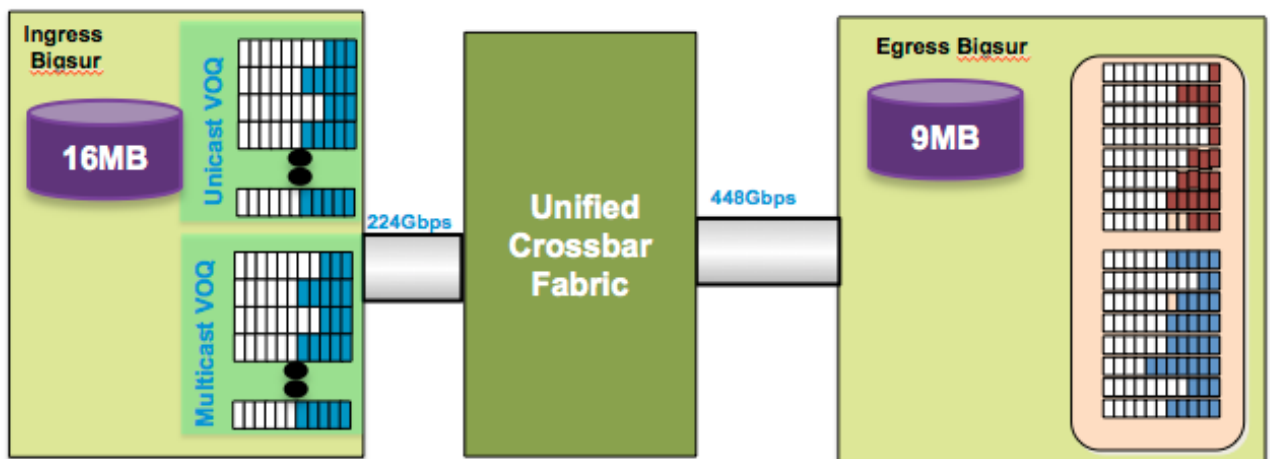
單播流量和緩衝

單點傳播流量先在輸出緩衝區池排隊，然後在輸出佇列已滿後在輸入緩衝區排隊，如下圖所示。



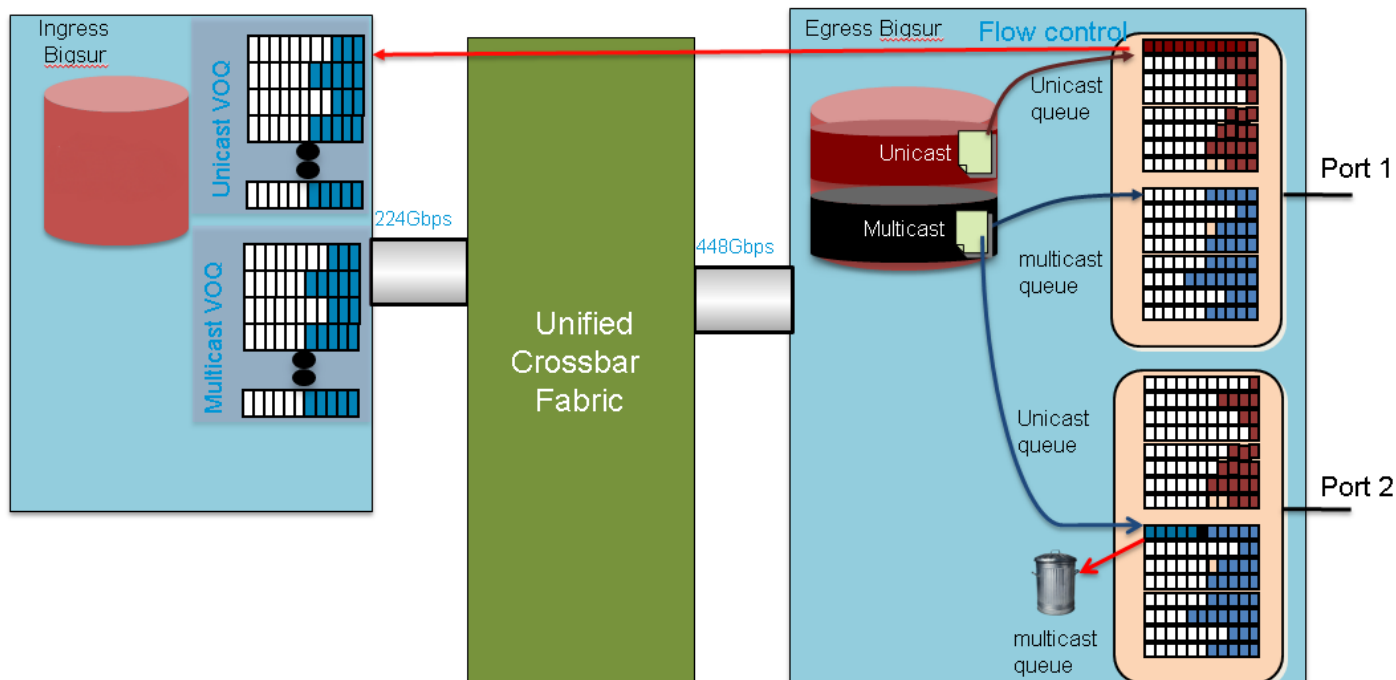
有16MB入口共用緩衝區和9MB出口共用緩衝區。緩衝區在12 x 10 gig埠或3 x 40 gig埠之間共用。共用緩衝區對突發吸收有好處。

以下是供您參考的記憶體分配直觀說明（Bigsur是ASIC/統一埠控制器的名稱），如下圖所示。



組播流量和緩衝

- 組播資料包在出口處緩衝和丟棄
- 在擁塞點附近捨棄多點傳送封包，以避免線路頭封鎖(HOLB)
- 維護單播的無損交換矩陣，如圖所示。



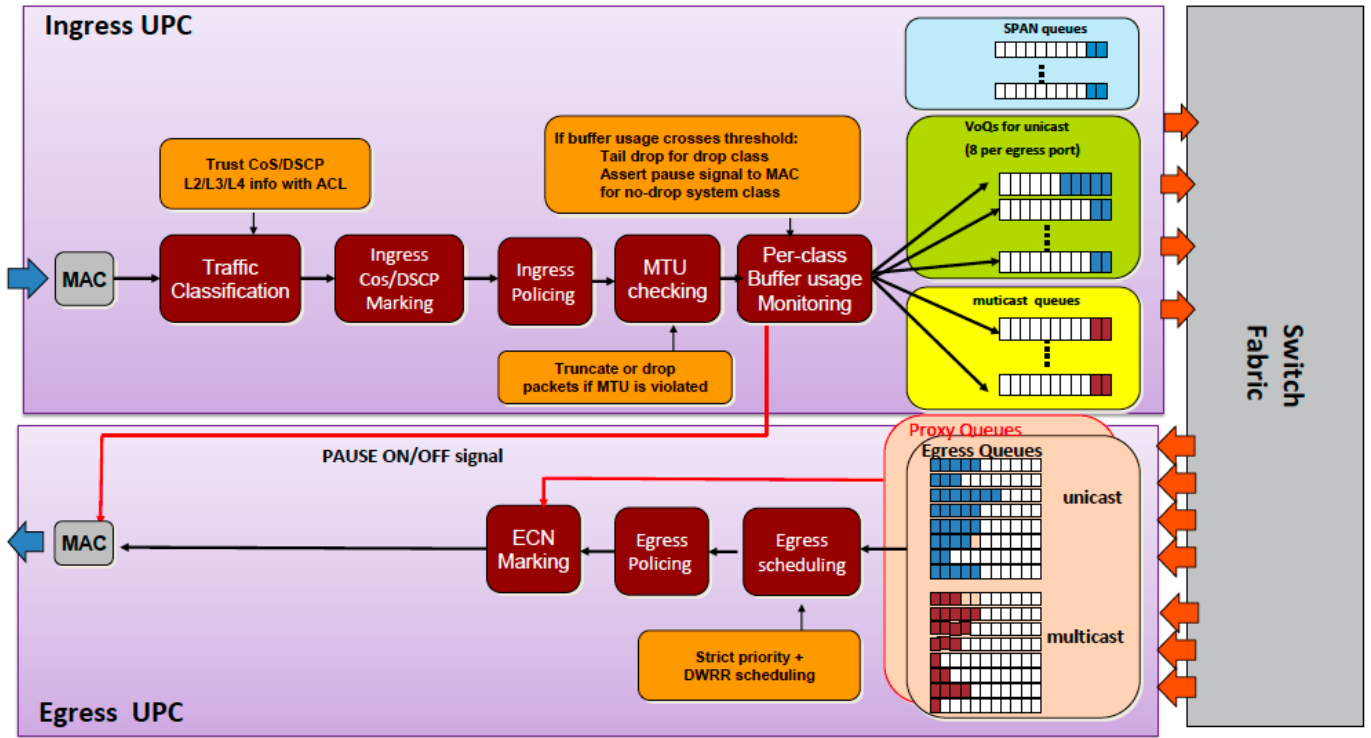
在大多數情況下，輸出丟棄始終是由於組播/廣播/未知的單點傳播流量。

什麼原因導致輸入丟棄？

擁塞的出口連線埠導致出口緩衝區先填滿，然後造成入口背壓。這僅適用於單播流量。輸入緩衝區滿後，您可能會捨棄輸入上的流量，從而導致輸入捨棄。

這種解釋的級別非常高，而且易於理解，但還有更多內容，尤其是當您檢視不同類別的流量、隊列等時。虛擬輸出佇列(VOQ)的概念在Nexus平台上經常使用。VOQ是每個輸出連線埠上每個IEEE 802.1p服務類別(CoS)的輸入緩衝區分配。因此每個出口埠有8個VOQ。

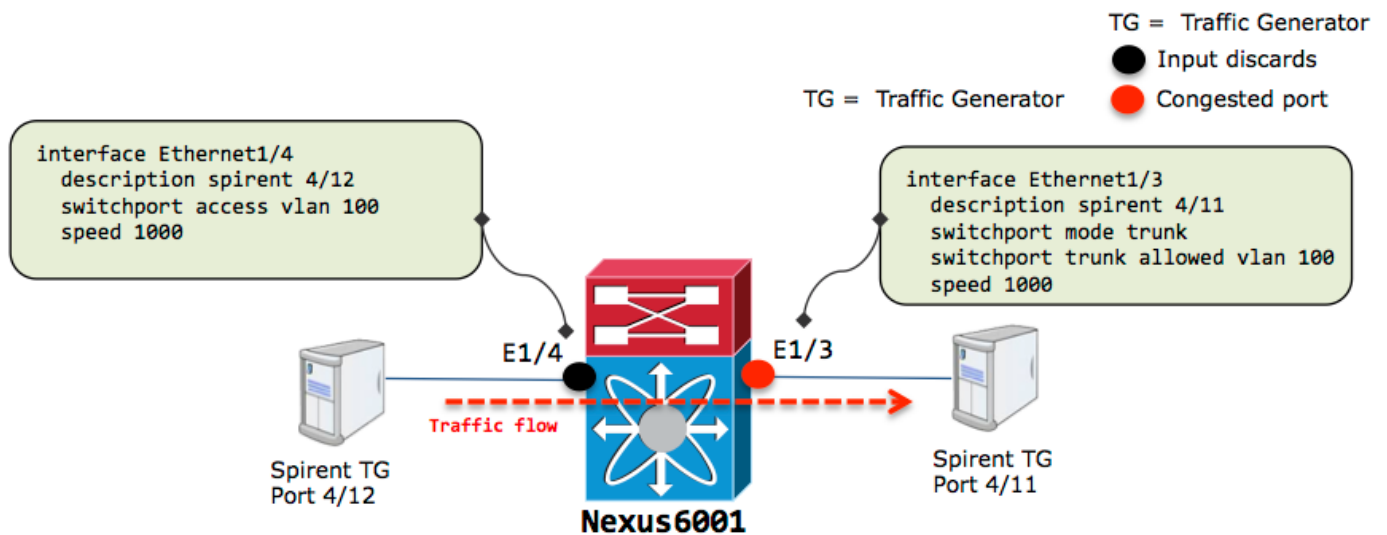
一個CoS中一個出口埠上的擁塞最終會滲入到其入口埠上相應VOQ的擁塞中。一旦達到限制，流量就會被丟棄。但是，它不影響發往其他CoS或其他出口介面的流量，從而避免HOLB，否則會引起擁塞擴散。從輸入連線埠到輸出連線埠和各個執行中的區塊的流量如下圖所示。



疑難排解案例

場景1.輸入丟棄

實驗室設定：



從e1/3開始並可能超訂用的線速流量：

```
nexus6001# sh int e1/3
Ethernet1/3 is up
Dedicated Interface
Hardware: 1000/10000 Ethernet, address: 002a.6a56.7a8a (bia 002a.6a56.7a8a)
Description: spirent 4/11
MTU 1500 bytes, BW 1000000 Kbit,, BW 1000000 Kbit, DLY 10 usec
reliability 255/255, txload 251/255, rxload 25/255
```

```
Encapsulation ARPA, medium is broadcast
Port mode is trunk
full-duplex, 1000 Mb/s
Beacon is turned off
Input flow-control is off, output flow-control is off
Switchport monitor is off
EtherType is 0x8100
Last link flapped 11:39:20
Last clearing of "show interface" counters 00:00:15
0 interface resets
30 seconds input rate 98683696 bits/sec, 8223 packets/sec
30 seconds output rate 986853640 bits/sec, 82019 packets/sec
Load-Interval #2: 5 minute (300 seconds)
  input rate 98.68 Mbps, 8.22 Kpps; output rate 986.85 Mbps, 82.01 Kpps
RX
```

```
124003 unicast packets  0 multicast packets  0 broadcast packets
124003 input packets  186004500 bytes
0 jumbo packets  0 storm suppression bytes
0 runts  0 giants  0 CRC  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

```
1236745 unicast packets  9 multicast packets  0 broadcast packets
1236754 output packets  1860065401 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
```

```
nexus6001# sh int e1/4
```

```
Ethernet1/4 is up
Dedicated Interface
```

```
Hardware: 1000/10000 Ethernet, address: 002a.6a56.7a8b (bia 002a.6a56.7a8b)
Description: spirent 4/12
MTU 1500 bytes, BW 1000000 Kbit,, BW 1000000 Kbit, DLY 10 usec
reliability 255/255, txload 25/255, rxload 251/255
Encapsulation ARPA, medium is broadcast
Port mode is access
full-duplex, 1000 Mb/s
Beacon is turned off
Input flow-control is off, output flow-control is off
Switchport monitor is off
EtherType is 0x8100
Last link flapped 10:53:31
Last clearing of "show interface" counters 00:00:04
0 interface resets
30 seconds input rate 986840376 bits/sec, 82236 packets/sec
30 seconds output rate 98421072 bits/sec, 8223 packets/sec
Load-Interval #2: 5 minute (300 seconds)
  input rate 986.84 Mbps, 82.23 Kpps; output rate 98.42 Mbps, 8.22 Kpps
RX
```

```
326332 unicast packets  0 multicast packets  0 broadcast packets
326332 input packets  489496500 bytes
0 jumbo packets  0 storm suppression bytes
0 runts  0 giants  0 CRC  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  863 input discard >>>>>
```

```

0 Rx pause
TX
32633 unicast packets 2 multicast packets 0 broadcast packets
32635 output packets 48819096 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

```

在模擬設定中（如此處），您知道超訂用的原因，但在流量量變曲線突發的生產設定中，通過這些命令找出擁塞的出口埠可能是一個挑戰。

此處列出的步驟可幫助您識別擁塞的出口埠。

步驟1. 使用輸入丟棄識別埠

在埠e1/4上看到輸入丟棄：

```

nexus6001# sh int e1/4 | in i disc
 0 input with dribble 3024 input discard
 0 lost carrier 0 no carrier 0 babble 0 output discard

```

```

nexus6001# sh queuing int e1/4
Ethernet1/4 queuing information:

```

```

TX Queuing
  qos-group  sched-type  oper-bandwidth
    0          WRR         100

```

```

RX Queuing

```

```

qos-group 0 >>>> Drops in QOS 0

```

```

q-size: 100160, q-size-40g: 100160, HW MTU: 1500 (1500 configured)

```

```

drop-type: drop, xon: 0, xoff: 0

```

```

Statistics:

```

```

Pkts received over the port           : 9612480
Ucast pkts sent to the cross-bar      : 9587016
Mcast pkts sent to the cross-bar      : 0
Ucast pkts received from the cross-bar : 961249
Pkts sent to the port                 : 961261
Pkts discarded on ingress             : 3024 >>>>>>
Per-priority-pause status             : Rx (Inactive), Tx (Inactive)

```

步驟2. ASIC標識

- 從此輸出將介面對映到內部ASIC(UPC)編號。
- 從您注意到捨棄的輸入連線埠ID中找出輸入ASIC ID。

```

nexus6001# sh hard internal bigsur all-ports

```

```

Bigsur Port Info:

```

Port name	asic idx	inst slot	inst asic	eport logi	flag adm	opr opr	if_index diag	ucVer
sup1	0 0	0 0	0 - 48	b3 en	dn 15010000	pass 0.00		
sup0	0 0	0 0	1 - 49	b3 en	dn 15020000	pass 0.00		
1gb1/1	1 0	1 1	2 - 0	b3 en	up 1a000000	pass 0.00		
1gb1/2	1 0	1 1	3 - 1	b3 en	up 1a001000	pass 0.00		
1gb1/3	1 0	1 1	0 - 2	b3 en	up 1a002000	pass 0.00		

```

1gb1/4 |1**|0|1|1-3|b3|en|up|1a003000|pass|0.00 >>>** is the asic number
1gb1/5 |1|0|1|6-4|b3|en|up|1a004000|pass|0.00
1gb1/6 |1|0|1|7-5|b3|en|up|1a005000|pass|0.00
1gb1/7 |1|0|1|4-6|b3|en|up|1a006000|pass|0.00
1gb1/8 |1|0|1|5-7|b3|en|up|1a007000|pass|0.00
1gb1/9 |1|0|1|10-8|b3|en|up|1a008000|pass|0.00
1gb1/10|1|0|1|11-9|b3|en|up|1a009000|pass|0.00
1gb1/11|1|0|1|8-10|b3|en|up|1a00a000|pass|0.00
xgb1/12|1|0|1|9-11|b3|en|dn|1a00b000|pass|0.00
xgb1/13|2|0|2|2-12|b3|en|dn|1a00c000|pass|0.00
xgb1/14|2|0|2|3-13|b3|en|dn|1a00d000|pass|0.00
xgb1/15|2|0|2|0-14|b3|en|dn|1a00e000|pass|0.00
xgb1/16|2|0|2|1-15|b3|en|dn|1a00f000|pass|0.00
xgb1/17|2|0|2|6-16|b3|en|dn|1a010000|pass|0.00
xgb1/18|2|0|2|7-17|b3|en|dn|1a011000|pass|0.00
xgb1/19|2|0|2|4-18|b3|en|dn|1a012000|pass|0.00
xgb1/20|2|0|2|5-19|b3|en|dn|1a013000|pass|0.00
xgb1/21|2|0|2|10-20|b3|en|dn|1a014000|pass|0.00
xgb1/22|2|0|2|11-21|b3|en|dn|1a015000|pass|0.00
xgb1/23|2|0|2|8-22|b3|en|dn|1a016000|pass|0.00
xgb1/24|2|0|2|9-23|b3|en|dn|1a017000|pass|0.00
xgb1/25|3|0|3|2-24|b3|en|dn|1a018000|pass|0.00
xgb1/26|3|0|3|3-25|b3|en|dn|1a019000|pass|0.00
xgb1/27|3|0|3|0-26|b3|en|dn|1a01a000|pass|0.00
xgb1/28|3|0|3|1-27|b3|en|dn|1a01b000|pass|0.00
xgb1/29|3|0|3|6-28|b3|en|dn|1a01c000|pass|0.00
xgb1/30|3|0|3|7-29|b3|en|dn|1a01d000|pass|0.00
xgb1/31|3|0|3|4-30|b3|en|dn|1a01e000|pass|0.00
xgb1/32|3|0|3|5-31|b3|en|dn|1a01f000|pass|0.00
xgb1/33|3|0|3|10-32|b3|en|dn|1a020000|pass|0.00
xgb1/34|3|0|3|11-33|b3|en|dn|1a021000|pass|0.00
xgb1/35|3|0|3|8-34|b3|en|dn|1a022000|pass|0.00
xgb1/36|3|0|3|9-35|b3|en|dn|1a023000|pass|0.00
xgb1/37|4|0|4|2-36|b3|en|dn|1a024000|pass|0.00
xgb1/38|4|0|4|3-37|b3|en|dn|1a025000|pass|0.00
xgb1/39|4|0|4|0-38|b3|en|dn|1a026000|pass|0.00
xgb1/40|4|0|4|1-39|b3|en|dn|1a027000|pass|0.00
xgb1/41|4|0|4|6-40|b3|en|dn|1a028000|pass|0.00
xgb1/42|4|0|4|7-41|b3|en|dn|1a029000|pass|0.00
xgb1/43|4|0|4|4-42|b3|en|dn|1a02a000|pass|0.00
xgb1/44|4|0|4|5-43|b3|en|dn|1a02b000|pass|0.00
xgb1/45|4|0|4|10-44|b3|en|dn|1a02c000|pass|0.00
xgb1/46|4|0|4|11-45|b3|en|dn|1a02d000|pass|0.00
xgb1/47|4|0|4|8-46|b3|en|dn|1a02e000|pass|0.00
xgb1/48|4|0|4|9-47|b3|en|dn|1a02f000|pass|0.00
40gb2/1|5|1|0|2-0|b3|dis|dn|1a0f0000|pass|0.00
40gb2/2|5|1|0|1-1|b3|dis|dn|1a0f1000|pass|0.00
40gb2/3|6|1|1|2-2|b3|dis|dn|1a0f2000|pass|0.00
40gb2/4|6|1|1|1-3|b3|dis|dn|1a0f3000|pass|0.00
Done.

```

步驟3.確定出口擁塞埠

- 使用VOQ計數器識別擁塞的出口埠。
- 使用counters voq asic-num中的ASIC編號可找出哪個出口埠導致丟包。

```

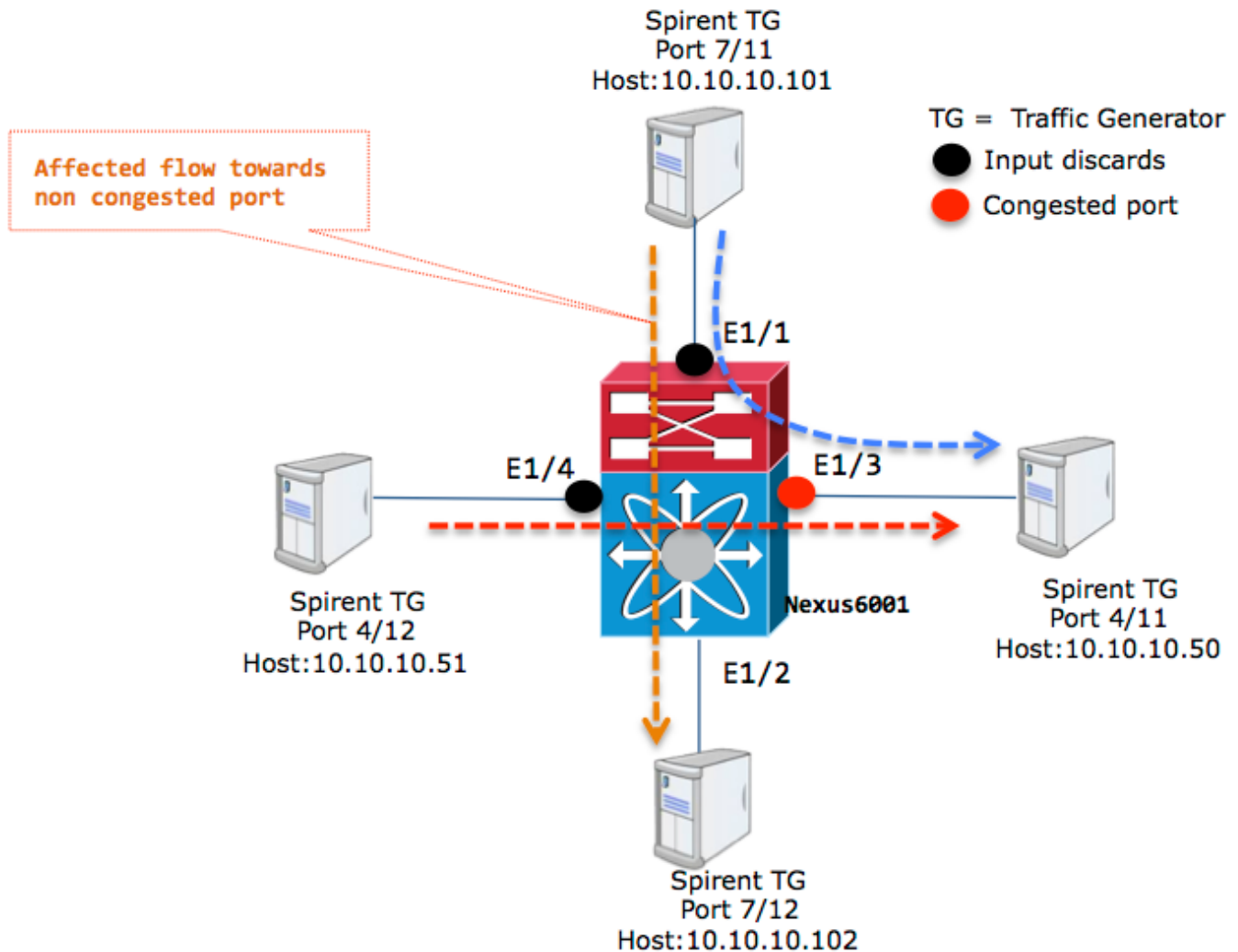
nexus6001# sh plat soft qd info counters voq asic-num 1
+-----+-----+-----+-----+
| port | TRANSMIT | TAIL DROP | HEAD DROP |
+-----+-----+-----+-----+
Eth1/3
  QUEUE-3          3222876464          8545008          0

```

Eth1/4				
QUEUE-3	323451170	0	0	
Eth1/6				
QUEUE-3	871362	0	0	
SUP_HI				
QUEUE-0	2041	0	0	

場景2.使用HOLB輸入丟棄

實驗室設定：



所有埠都位於vlan 100中。

您可以看到e1/4和e1/1上的輸入捨棄器，這取決於流向主機1/3的輸入介面的流量速率。

```
nexus6001# sh int e1/4 | in discard|rate
30 seconds input rate 592103840 bits/sec, 49341 packets/sec
30 seconds output rate 166412120 bits/sec, 13863 packets/sec
input rate 592.10 Mbps, 49.34 Kpps; output rate 834.82 Mbps, 69.55 Kpps
0 input with dribble 15245 input discard
0 lost carrier 0 no carrier 0 babble 0 output discard
```

```
nexus6001# sh int e1/1 | in discard|rate
```



```

30 seconds input rate 986839872 bits/sec, 82236 packets/sec
30 seconds output rate 99790992 bits/sec, 8310 packets/sec
input rate 986.84 Mbps, 82.23 Kpps; output rate 500.88 Mbps, 41.73 Kpps
0 input with dribble 110632 input discard
0 lost carrier 0 no carrier 0 babble 0 output discard

```

使用與場景1相同的過程。您可以找到出口擁塞埠。

```

nexus6001# sh plat so qd info counters voq ASIC-num 1 <snip>
+-----+-----+-----+-----+-----+
| port | TRANSMIT | TAIL DROP | HEAD DROP |
+-----+-----+-----+-----+
Eth1/3
QUEUE-3 3893719464 164782171 0

```

必須受影響的流是指向10.10.10.50。10.10.101和10.10.10.102之間的流必須乾淨。

但事實並非如此。停滯或緩慢排出的出口埠可能會導致一個或多個向出口埠傳送流量的輸入埠上的所有緩衝區耗盡，從而影響這些輸入埠上的所有流量。這是典型的HOLB問題。

Spirent流量生成器顯示流量被丟棄。埠號為Spirent埠號，如下圖所示。

Name/ID	Tx Port Name	Rx Port Names	Tx Count (Frames)	Rx Count (Frames)	Dropped Count (Frames)	Dropped Frame Percent	In-order Count (Frames)	Reordered Count (Frames)
StreamBloc...	Port //4/11	Port //4/12	0	0	0	0.000	0	0
StreamBloc...	Port //4/12	N/A	0	0	0	0.000	0	0
StreamBloc...	Port //4/12	Port //4/11	1,307,568	1,100,070	223,516	16.887	1,100,070	0
StreamBloc...	Port //7/11	Port //7/12	461,229	275,398	172,495	38.512	275,398	0
StreamBloc...	Port //7/11	Port //4/11	1,844,950	1,100,058	664,699	37.665	1,100,058	0

HOLB緩解：啟用VOQ限制

為了避免這種情況，可以使用設定的閾值配置VOQ（僅適用於單播流量）。

```

nexus6001(config)# hard unicast voq-limit

```

設定後，流向非擁塞連線埠的流量不會受到影響。

VOQ限制配置之後的思波流量發生器檢視如下圖所示。

Name/ID	Tx Port Name	Rx Port Names	Tx Count (Frames)	Rx Count (Frames)	Dropped Count (Frames)	Dropped Frame Percent	In-order Count (Frames)	Reordered Count (Frames)
StreamBloc...	Port //4/11	Port //4/12	0	0	0	0.000	0	0
StreamBloc...	Port //4/12	N/A	0	0	0	0.000	0	0
StreamBloc...	Port //4/12	Port //4/11	1,348,359	1,133,953	230,398	16.887	1,133,953	0
StreamBloc...	Port //7/11	Port //7/12	474,821	461,488	0	0.000	461,488	0
StreamBloc...	Port //7/11	Port //4/11	1,899,318	1,133,940	685,182	37.665	1,133,940	0

雖然此組態顯示明顯的優勢，可防止由於HOLB而發生捨棄。為什麼這不是預設配置？

通常，生產環境中的流量可能會在自然界中爆發。通過禁用VOQ閾值，您允許入口緩衝區吸收流量微突發，而無需丟棄。

除非這種情況要求必須啟用VOQ限制，否則建議使用預設值，即將其禁用。

HOLB緩解：流量分類

還有一種方法可以利用QoS配置來緩解HOLB。由於入口丟棄僅影響特定的VOQ (VOQ又是一個特定的QoS類)，因此您可以將受影響的流量對映到非擁塞埠到不同的QoS組。從此輸出中，入口丟棄會影響QoS組0類。

```
nexus6001# sh queuing int e1/4
Ethernet1/4 queuing information:
TX Queuing
  qos-group  sched-type  oper-bandwidth
    0          WRR        100

RX Queuing
qos-group 0 >>>> Drops in QoS 0
q-size: 100160, q-size-40g: 100160, HW MTU: 1500 (1500 configured)
drop-type: drop, xon: 0, xoff: 0
Statistics:
  Pkts received over the port          : 9612480
  Ucast pkts sent to the cross-bar     : 9587016
  Mcast pkts sent to the cross-bar     : 0
  Ucast pkts received from the cross-bar : 961249
  Pkts sent to the port                : 961261
Pkts discarded on ingress            : 3024 >>>>>
  Per-priority-pause status           : Rx (Inactive), Tx (Inactive)
```

此處的配置將重要流量對映到QoS組2。

1. 為不能丟棄的流量定義ACL。目標是將此流量分類為不同的QoS組，使其不會受到影響。

```
ip access-list SINGLEFLOW
  statistics per-entry
  10 permit ip 10.10.10.101/32 10.10.10.102/32
```

2. QoS分類：

```
class-map type qos match-all FIX_AFFECTED_FLOW
  match access-group name SINGLEFLOW
policy-map type qos QOS_POLICY_FIX_AFFECTED_FLOW
  class FIX_AFFECTED_FLOW
    set qos-group 2
```

3. 網路QoS配置：

```
class-map type network-qos QOSGRP2
  match qos-group 2
policy-map type network-qos NQOS-GRP2
  class type network-qos QOSGRP2
  class type network-qos class-default
```

4. 應用各種政策。網路的QoS是系統範圍的，而分類策略可以應用於單個介面。

```
system qos
service-policy type network-qos NQOS-GRP2

interface Ethernet1/1
service-policy type qos input QOS_POLICY_FIX_AFFECTED_FLOW
```

5. QoS組2類沒有發現丟棄：

```
nexus6001(config-if)# sh queuing int e1/1
Ethernet1/1 queuing information:
TX Queuing
  qos-group  sched-type  oper-bandwidth
    0         WRR        100
    2         WRR        0
RX Queuing
  qos-group 0
  q-size: 100160, q-size-40g: 100160, HW MTU: 1500 (1500 configured)
  drop-type: drop, xon: 0, xoff: 0
  Statistics:
    Pkts received over the port          : 525111
    Ucast pkts sent to the cross-bar     : 327510
    Mcast pkts sent to the cross-bar     : 0
    Ucast pkts received from the cross-bar : 0
    Pkts sent to the port                 : 0
    Pkts discarded on ingress           : 197868 >>>>
    Per-priority-pause status            : Rx (Inactive), Tx (Inactive)
  qos-group 2
  q-size: 100160, q-size-40g: 100160, HW MTU: 1500 (1500 configured)
  drop-type: drop, xon: 0, xoff: 0
  Statistics:
    Pkts received over the port          : 131413
    Ucast pkts sent to the cross-bar     : 132096
    Mcast pkts sent to the cross-bar     : 0
    Ucast pkts received from the cross-bar : 0
    Pkts sent to the port                 : 0
    Pkts discarded on ingress           : 0 >>> No Drops
    Per-priority-pause status            : Rx (Inactive), Tx (Inactive)
```

相關資訊

- [Nexus 6000系列交換機QoS配置示例](#)
- [技術支援與文件 - Cisco Systems](#)