

Nexus 3500輸出捨棄和緩衝區QoS

目錄

[簡介](#)

[方法](#)

[檢查輸出丟棄](#)

[確定丟棄是單播還是組播](#)

[確定使用的輸出緩衝區](#)

[檢查活動緩衝區監控](#)

[計數器主動增加](#)

[簡要輸出](#)

[詳細輸出](#)

[超過閾值時生成日誌](#)

[值得注意的思科錯誤ID](#)

[常見問題](#)

[附錄 — 功能資訊](#)

[緩衝區管理](#)

[計畫](#)

[多點傳送慢速接收器](#)

[主動緩衝區監控](#)

[硬體實施](#)

[軟體實施](#)

簡介

本檔案介紹用於對Nexus 3500平台上捨棄的流量型別以及捨棄此流量的輸出緩衝區(OB)進行疑難排解的命令。

方法

1. [檢查輸出丟棄](#)
2. [確定丟棄是單播還是組播](#)
3. [確定使用的輸出緩衝區](#)
4. [檢查活動緩衝區監控](#)

檢查輸出丟棄

檢查實體介面統計資料，以確定流量是否已在輸出方向捨棄。確定TX方向的「輸出丟棄」計數器是否遞增和/或非零。

```
Nexus3548# show interface Eth1/7
Ethernet1/7 is up
  Dedicated Interface
  Hardware: 100/1000/10000 Ethernet, address: a44c.116a.913c (bia a44c.116a.91ee)
```

```

Description: Unicast Only
Internet Address is 1.2.1.13/30
MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec
reliability 255/255, txload 35/255, rxload 1/255
Encapsulation ARPA
full-duplex, 1000 Mb/s, media type is 1G
Beacon is turned off
Input flow-control is off, output flow-control is off
Rate mode is dedicated
Switchport monitor is off
EtherType is 0x8100
Last link flapped 00:03:48
Last clearing of "show interface" counters 00:03:55
1 interface resets
30 seconds input rate 200 bits/sec, 0 packets/sec
30 seconds output rate 0 bits/sec, 0 packets/sec
Load-Interval #2: 5 minute (300 seconds)
  input rate 40 bps, 0 pps; output rate 139.46 Mbps, 136.16 Kpps
RX
  1 unicast packets  118 multicast packets  0 broadcast packets
  119 input packets  9830 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
  23605277 unicast packets  0 multicast packets  0 broadcast packets
  23605277 output packets  3038908385 bytes
  0 jumbo packets
  0 output errors  0 collision  0 deferred  0 late collision
  0 lost carrier  0 no carrier  0 babble 11712542 output discard
  0 Tx pause

```

確定丟棄是單播還是組播

一旦確定介面丟棄流量，請輸入 `show queuing interface <x/y>` 命令以查明丟棄的流量是組播還是單播。在低於 6.0(2)A3(1) 的版本中，輸出如下所示：

```

Nexus3548# show queuing interface Eth1/7
Ethernet1/7 queuing information:
TX Queuing
  qos-group  sched-type  oper-bandwidth
  0          WRR        100

RX Queuing
Multicast statistics:
  Mcast pkts dropped          : 0
Unicast statistics:
  qos-group 0
  HW MTU: 1500 (1500 configured)
  drop-type: drop, xon: 0, xoff: 0
Statistics:
  Ucast pkts dropped          : 11712542

```

在版本 6.0(2)A3(1) 及更高版本中，輸出如下所示：

```

Nexus3548# show queuing interface Eth1/7
Ethernet1/7 queuing information:
  qos-group  sched-type  oper-bandwidth

```

```

0          WRR          100
Multicast statistics:
  Mcast pkts dropped          : 0
Unicast statistics:
  qos-group 0
  HW MTU: 1500 (1500 configured)
  drop-type: drop, xon: 0, xoff: 0
  Statistics:
Ucast pkts dropped          : 11712542

```

附註：如果為連線埠設定了多點傳送慢速接收器，請參閱獲得功能資訊，因為硬體限制，所以不會使用show queuing interface Eth<x/y>命令追蹤捨棄。請參閱思科錯誤ID [CSCuj2106](#)。

確定使用的輸出緩衝區在Nexus 3500中，有三個緩衝池用於輸出方向。show hardware internal mtc-usd info port-mapping命令的輸出提供了對映資訊。

```
Nexus3548# show hardware internal mtc-usd info port-mapping
```

```
OB Ports to Front Ports:
```

```

===== OB0 =====      ===== OB1 =====      ===== OB2 =====
45 47 21 23 09 11 33 35    17 19 05 07 41 43 29 31    13 15 37 39 25 27 01 03
46 48 22 24 10 12 34 36    18 20 06 08 42 44 30 32    14 16 38 40 26 28 02 04

```

```
Front Ports to OB Ports:
```

```

=OB2= =OB1= =OB0= =OB2=    =OB1= =OB0= =OB2= =OB1=    =OB0= =OB2= =OB1= =OB0=
12 14 04 06 08 10 00 02    00 02 04 06 08 10 12 14    12 14 04 06 08 10 00 02
13 15 05 07 09 11 01 03    01 03 05 07 09 11 13 15    13 15 05 07 09 11 01 03

```

```
Front port numbering (i.e. "01" here is e1/1):
```

```

=OB2= =OB1= =OB0= =OB2=    =OB1= =OB0= =OB2= =OB1=    =OB0= =OB2= =OB1= =OB0= 01 03 05 07 09 11
13 15    17 19 21 23 25 27 29 31    33 35 37 39 41 43 45 47 02 04 06 08 10 12 14 16    18 20 22
24 26 28 30 32    34 36 38 40 42 44 46 48

```

Note: Text in Red font is not CLI output, it's purely to help those reading the document faster match the actual front port instead of having to manually count up.

結果的第一部分表明OB池0用於前埠45、46、47、48等，OB1用於前埠17、18等。第二部分結果表明，Eth1/1對映到OB2埠12，Eth1/2對映到OB2埠13，等等。討論的埠Eth1/7對映到OB1。有關詳細資訊，請參閱本文檔中的**緩衝區管理**部分。檢查活動緩衝區監控有關此功能的詳細資訊，請參閱[Cisco Nexus 3548活動緩衝區監控白皮書](#)和本文檔中的部分。計數器主動增加如果輸出丟棄主動遞增，請使用此命令啟用主動緩衝區監視(ABM)。請注意，此命令允許您監控單播或組播，但不能同時監控兩者。此外，它還允許您配置取樣間隔和閾值。

```
hardware profile buffer monitor [unicast|multicast] {[sampling
```

簡要輸出啟用ABM後，您可以使用此命令檢視結果。

```
Nexus3500# show hardware profile buffer monitor interface e1/7 brief
```

```
Brief CLI issued at: 09/30/2013 19:43:50
```

```

          Maximum buffer utilization detected
          1sec      5sec      60sec      5min      1hr
          -----
Ethernet1/7    5376KB    5376KB    5376KB      N/A      N/A

```

這些結果表明，過去60秒內離開Eth1/7的單點傳播流量使用了OB1緩衝區的6 MB中的5.376 MB。

詳細輸出

```
Nexus3500# show hardware profile buffer monitor interface Eth1/7 detail
```

```
Detail CLI issued at: 09/30/2013 19:47:01
```

Legend -

- 384KB - between 1 and 384KB of shared buffer consumed by port
- 768KB - between 385 and 768KB of shared buffer consumed by port
- 307us - estimated max time to drain the buffer at 10Gbps

Active Buffer Monitoring for Ethernet1/7 is: Active

```

KBytes          384  768 1152 1536 1920 2304 2688 3072 3456 3840 4224 4608 4992 5376 5760
6144
us @ 10Gbps     307  614  921 1228 1535 1842 2149 2456 2763 3070 3377 3684 3991 4298 4605
4912

```

```

-----
-----
09/30/2013 19:47:01      0   0   0   0   0   0   0   0   0   0   0   0   0   0   250
0   0
09/30/2013 19:47:00      0   0   0   0   0   0   0   0   0   0   0   0   0   0   252
0   0
09/30/2013 19:46:59      0   0   0   0   0   0   0   0   0   0   0   0   0   0   253
0   0
09/30/2013 19:46:58      0   0   0   0   0   0   0   0   0   0   0   0   0   0   250
0   0
09/30/2013 19:46:57      0   0   0   0   0   0   0   0   0   0   0   0   0   0   250
0   0
09/30/2013 19:46:56      0   0   0   0   0   0   0   0   0   0   0   0   0   0   250
0   0
09/30/2013 19:46:55      0   0   0   0   0   0   0   0   0   0   0   0   0   0   251
0   0
09/30/2013 19:46:54      0   0   0   0   0   0   0   0   0   0   0   0   0   0   251
0   0
09/30/2013 19:46:53      0   0   0   0   0   0   0   0   0   0   0   0   0   0   250
0   0
09/30/2013 19:46:52      0   0   0   0   0   0   0   0   0   0   0   0   0   0   253
0   0
09/30/2013 19:46:51      0   0   0   0   0   0   0   0   0   0   0   0   0   0   249
0   0
...

```

每一行中的資訊以第二個間隔記錄。每列表示緩衝區使用情況。如命令結果中所述，如果為列「384」報告了一個非零值，這意味著當ABM輪詢該OB使用情況時，緩衝區使用率介於0-384 KB之間。Non-zero number是報告的使用次數。這些結果表明，在Eth1/7的最後10秒內，OB1的平均使用量為每秒249 - 253次，介於5.376 MB之間。清除此流量的緩衝區需要4298微秒(us)。超過閾值時生成日誌如果丟棄計數器和緩衝區使用率定期增加，則可以在超過閾值時設定閾值並生成日誌消息。

logging level mtc-usd 5

hardware profile buffer monitor unicast sampling 10 threshold 4608

該命令設定為以10納秒時間間隔監控單播流量，當流量超過緩衝區的75%時，它會生成日誌。您還可以建立排程程式，以便每小時收集ABM統計資訊和介面計數器輸出，並將其附加到bootflash檔案。以下範例適用於多點傳播流量：

hardware profile buffer monitor multicast

feature scheduler

scheduler job name ABM

show hardware profile buffer monitor detail >> ABMDetail.txt

show clock >> ABMBrief.txt

show hardware profile buffer monitor brief >> ABMBrief.txt

show clock >> InterfaceCounters.txt

show interface counters errors >> InterfaceCounters.txt

scheduler schedule name ABM

time start now repeat 1:0

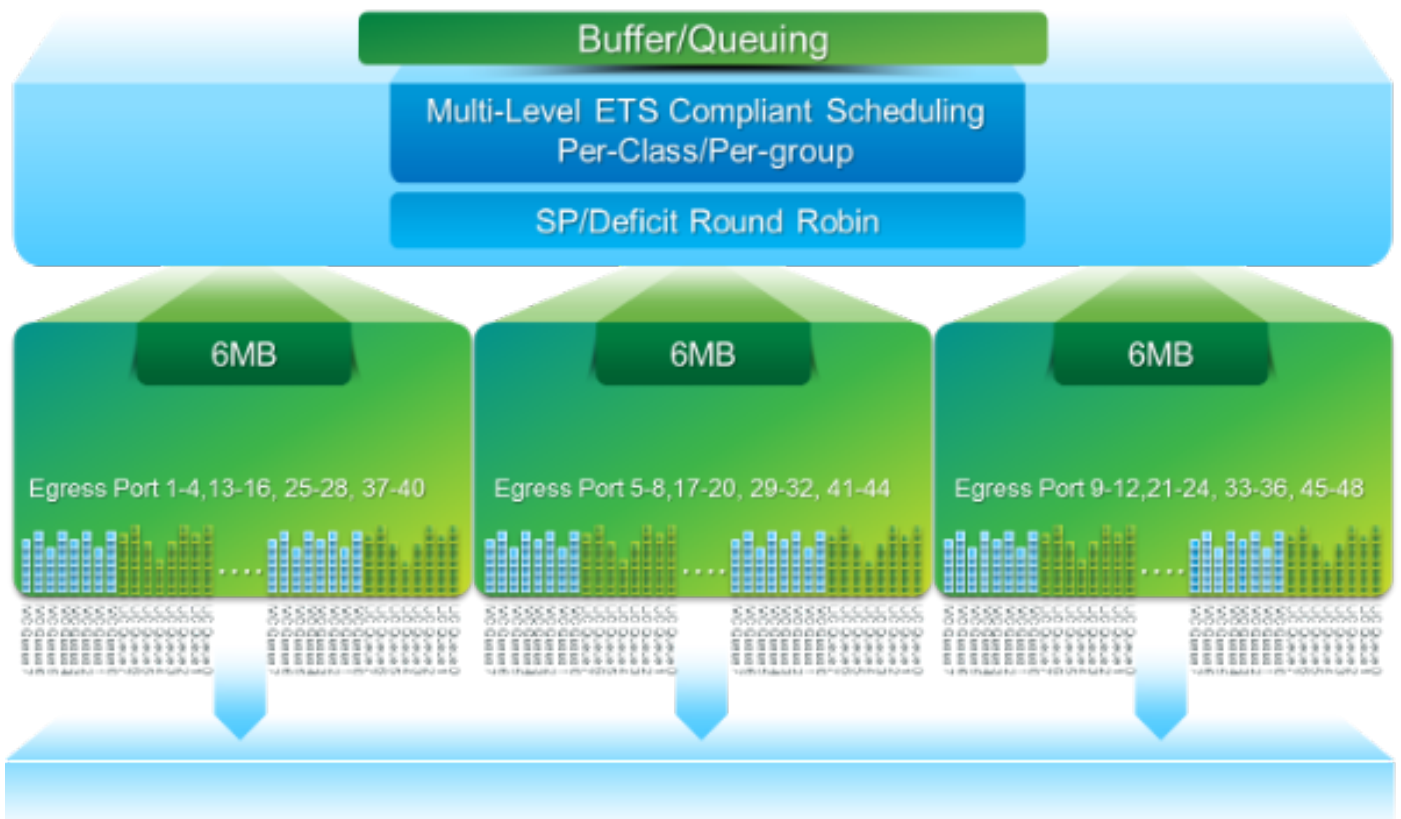
job name ABM

值得注意的思科錯誤ID

- 思科錯誤ID [CSCum21350](#):快速埠擺動會導致同一QoS緩衝區中的所有埠丟棄所有TX組播/廣播流量。版本6.0(2)A1(1d)及更新版本已修復此問題。

- 思科錯誤ID [CSCuq96923](#):多點傳送緩衝區區塊堵塞，導致輸出多點傳送/廣播捨棄。這一問題仍在調查中。
- 思科錯誤ID [CSCva20344](#):Nexus 3500緩衝區塊/鎖定 — 無TX多點傳送或廣播。不可複製的問題，可能會在版本6.0(2)U6(7)、6.0(2)A6(8)和6.0(2)A8(3)中得到修復。
- 思科錯誤ID [CSCvi93997](#): Cisco Nexus 3500交換機輸出緩衝區塊卡住。已在版本7.0(3)I7(8)和9.3(3)中修正。

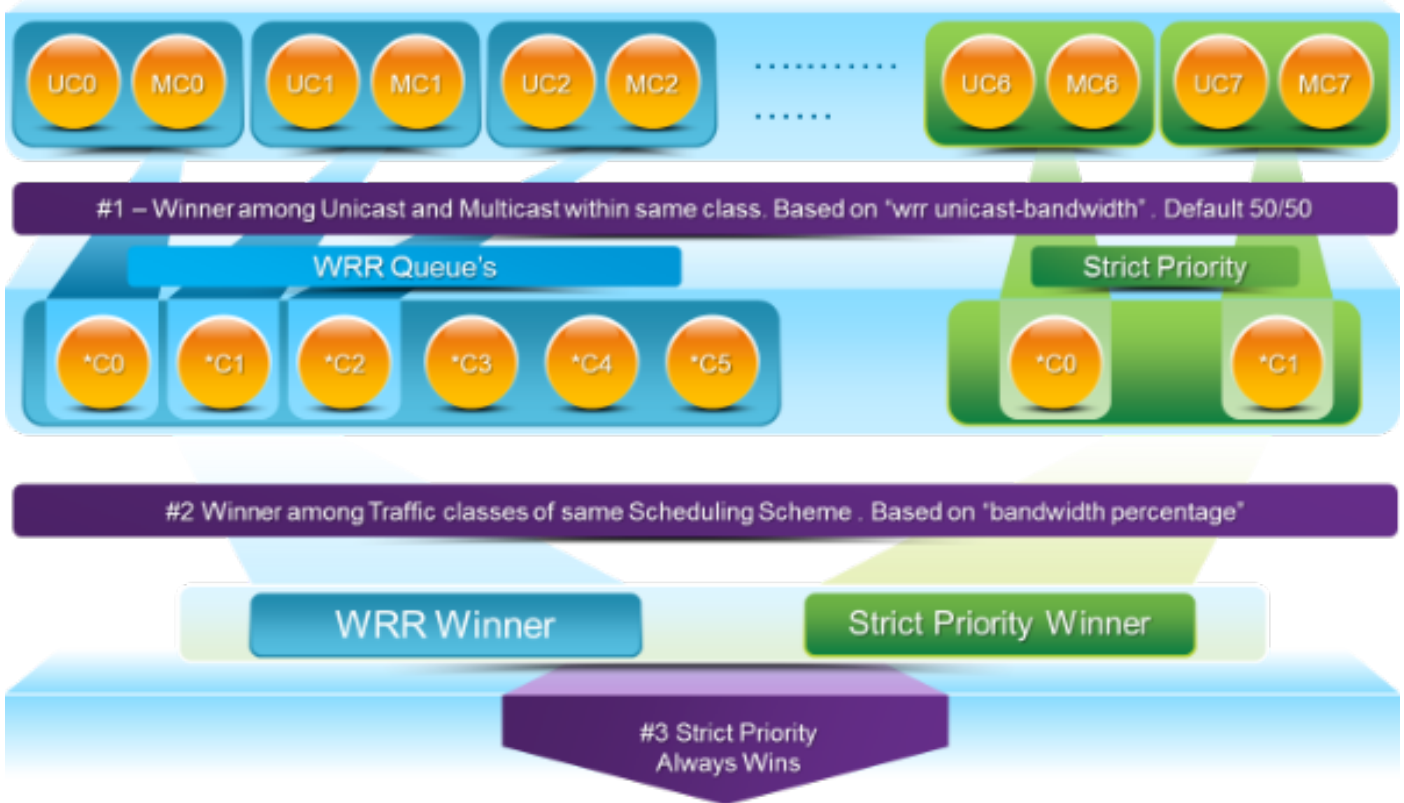
常見問題ABM是否會影響效能或延遲？不，此功能不會影響裝置的延遲或效能。較低的ABM硬體輪詢間隔會產生什麼影響？預設情況下，硬體輪詢間隔為4毫秒。您可以將此值配置為10納秒。由於硬體輪詢間隔較低，因此對效能或延遲沒有影響。選擇預設硬體輪詢4毫秒，以確保在軟體每秒鐘輪詢之前不會使直方圖計數器溢位。如果降低硬體輪詢間隔，則硬體計數器可能會達到255個取樣。由於CPU和記憶體限制，裝置無法處理低於一秒的軟體輪詢，以便匹配較低的硬體輪詢。該白皮書提供了較低硬體輪詢間隔的示例及其使用案例。**附錄 — 功能資訊緩衝區管理**



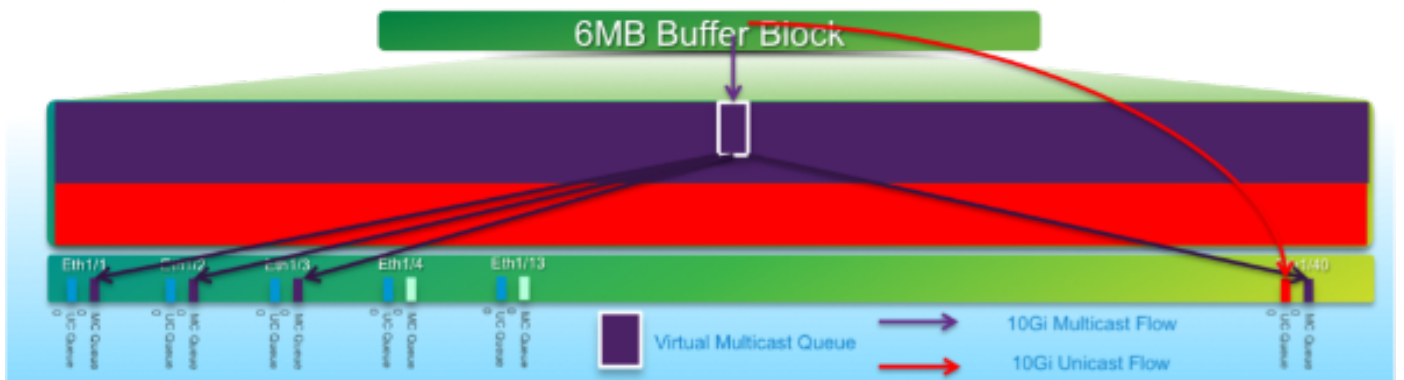
- 三個OB塊共用的18 MB資料包緩衝區：保留約4 MB:基於配置的最大傳輸單元(MTU)的大小 (每埠總和2 x MTU大小x已啟用QoS組的數量) 約14 MB共用：總緩衝區的剩餘部分大約767 KB的OB:0，用於CPU目標資料包
- 每個OB的6 MB由一組16個埠共用(show hardware internal mtc-usd info port-mapping命令)

計畫三層排程：

- 單播和組播
- 相同排程方案的流量類
- 方案中的流量類



多點傳送慢速接收器



在此圖中：

- 1 G Eth1/40上引入持續擁塞。
- 緩衝區塊上的其他多點傳送接收器(Eth1/1 - 3)會因多點傳送排程行為而受到影響。其他緩衝塊上的接收器不受影響。
- 「Multicast slow-receiver」可以應用於e1/40，以避免非擁塞連線埠上的流量損失。
- 在Eth1/40上，「Multicast slow-receiver」以10 G的速率耗盡組播。預計在擁塞的埠上仍會發生丟棄。
- 已使用hardware profile multicast slow-receiver port <x>命令進行配置。

主動緩衝區監控有關此功能的概述，請參閱[Cisco Nexus 3548活動緩衝器監控白皮書](#)。硬體實施

- ASIC有18個儲存桶，每個儲存桶對應於一個緩衝區利用率範圍（例如0-384KB、385-768KB等）。
- ASIC每4毫秒（預設）輪詢一次所有埠的緩衝區利用率。此ASIC輪詢間隔可配置為10納秒。
- 基於每個硬體輪詢間隔的緩衝區利用率，相應範圍的桶計數器會增加。也就是說，如果埠25消耗500 KB的緩衝區，則桶計數#2(385-768KB)會增加。
- 為直方圖格式的每個介面維護此緩衝區利用率計數器。
- 每個儲存桶用8位表示，因此計數器在255處最大，並且一旦軟體讀取資料就會重置計數器。

軟體實施

- 軟體每秒輪詢ASIC，以便下載並清除所有直方圖計數器。
- 這些直方圖計數器在儲存器中以一秒粒度保持60分鐘。
- 軟體還確保每小時將緩衝區直方圖複製到bootflash中，然後將這些直方圖複製到分析器中進行進一步分析。
- 這有效地維護了所有埠兩個小時的緩衝直方圖資料，最晚一個小時位於記憶體中，第二個小時位於bootflash中。