

Caídas de salida de Nexus 3500 y QoS de búfer

Contenido

[Introducción](#)

[Metodología](#)

[Comprobar caídas de salida](#)

[Determinar si las caídas son unidifusión o multidifusión](#)

[Determinar qué búfer de salida se utiliza](#)

[Verifique la Supervisión del búfer activo](#)

[Los contadores aumentan activamente](#)

[Breve salida](#)

[Salida detallada](#)

[Generar un registro cuando se cruza un umbral](#)

[ID de bug de Cisco notables](#)

[Preguntas Frecuentes](#)

[Apéndice - Información sobre funciones](#)

[Administración de búfer](#)

[Planificación](#)

[Receptor lento de multidifusión](#)

[Supervisión de búfer activa](#)

[Implementación de hardware](#)

[Implementación de software](#)

Introducción

Este documento describe los comandos utilizados para resolver problemas del tipo de tráfico descartado en la plataforma Nexus 3500 y el búfer de salida (OB) en el que se descarta este tráfico.

Metodología

1. [Comprobar caídas de salida](#)
2. [Determinar si las caídas son unidifusión o multidifusión](#)
3. [Determinar qué búfer de salida se utiliza](#)
4. [Verifique la Supervisión del búfer activo](#)

Comprobar caídas de salida

Verifique las estadísticas de la interfaz física para determinar si el tráfico se descarta en la dirección de salida. Determine si el contador "descarte de salida" en la dirección TX aumenta y/o es distinto de cero.

```

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

```

Determinar si las caídas son unidifusión o multidifusión

Una vez que se determina que la interfaz descarta el tráfico, ingrese el comando **show queuing interface <x/y>** para averiguar si el tráfico descartado es multicast o unicast. En las versiones anteriores a 6.0(2)A3(1), la salida es similar a:

```

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

```

En la versión 6.0(2)A3(1) y posteriores, la salida es similar a:

```

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

```

Nota: Si el receptor lento multicast está configurado para el puerto, consulte para obtener información sobre la función, las caídas no se rastrean con el comando `show queuing interface Eth<x/y>` debido a una limitación de hardware. Consulte Cisco bug ID [CSCuj21006](https://www.cisco.com/cisco/webbugtool/show_bug.do?bugID=CSCuj21006).

Determinar qué búfer de salida se utiliza

En Nexus 3500, se utilizan tres conjuntos de búfer en la dirección de salida. La salida del comando `show hardware internal mtc-usd info port-mapping` proporciona la información de mapping.

```

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.

La primera parte de los resultados indica que los puertos frontales como 45, 46, 47, 48, etc. utilizan el grupo OB 0 y OB1 los puertos frontales 17, 18, etc.

La segunda parte de los resultados indica que Eth1/1 se mapea al puerto OB2 12, Eth1/2 se mapea al puerto OB2 13, y así sucesivamente.

El puerto en discusión, Eth1/7, se mapea a OB1.

Consulte la sección [Administración de Búfer](#) en este documento para obtener más información.

Verifique la Supervisión del búfer activo


```

0      0
09/30/2013 19:46:57      0      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      0      250
0      0
09/30/2013 19:46:55      0      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      0      251
0      0
09/30/2013 19:46:53      0      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      0      253
0      0
09/30/2013 19:46:51      0      0      0      0      0      0      0      0      0      0      0      0      0      0      0      249
0      0
...

```

La información de cada fila se registra en un segundo intervalo. Cada columna representa el uso del búfer. Como se menciona en los resultados del comando, si se informa de un valor distinto de cero para la columna "384", significa que el uso del búfer estaba entre 0-384 KBytes cuando el ABM sondeó el uso de OB. El número distinto de cero es el número de veces que se informó del uso.

Estos resultados indican que el OB1 promedió 5,376 MB de uso entre 249 y 253 veces por segundo en los últimos 10 segundos para Eth1/7. Se necesitan 4298 microsegundos (us) para despejar el búfer de este tráfico.

Generar un registro cuando se cruza un umbral

Si el contador de caídas y el uso del búfer aumentan periódicamente, es posible establecer un umbral y generar un mensaje de registro cuando se cruza el umbral.

```

logging level mtc-usd 5
hardware profile buffer monitor unicast sampling 10 threshold 4608

```

El comando se configura para monitorear el tráfico de unidifusión en un intervalo de 10 nanosegundos y cuando supera el 75% del buffer genera un registro.

También puede crear un planificador para recopilar las estadísticas de ABM y la salida del contador de interfaz cada hora y anexarla a los archivos bootflash. Este ejemplo es para el tráfico multicast:

```

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

```

ID de bug de Cisco notables

- Id. de bug Cisco [CSCum21350](#): Los vaciados rápidos de puertos hacen que todos los puertos en el mismo búfer de QoS descarten todo el tráfico de multidifusión/difusión TX. Esto se corrige en la versión 6.0(2)A1(1d) y posteriores.
- Id. de bug Cisco [CSCuq96923](#): El bloque de búfer multicast está atascado, lo que da como resultado caídas de multicast/broadcast de salida. Esta cuestión sigue siendo objeto de investigación.
- Id. de bug Cisco [CSCva20344](#): Bloqueo/bloqueo del búfer Nexus 3500: sin multidifusión TX ni difusión. Problema no reproducible, potencialmente corregido en las versiones 6.0(2)U6(7), 6.0(2)A6(8) y 6.0(2)A8(3).
- Id. de error de Cisco [CSCvi93997](#): Bloque del búfer de salida de switches Cisco Nexus 3500 atascado. Esto se fija en las versiones 7.0(3)I7(8) y 9.3(3).

Preguntas Frecuentes

¿El ABM afecta al rendimiento o a la latencia?

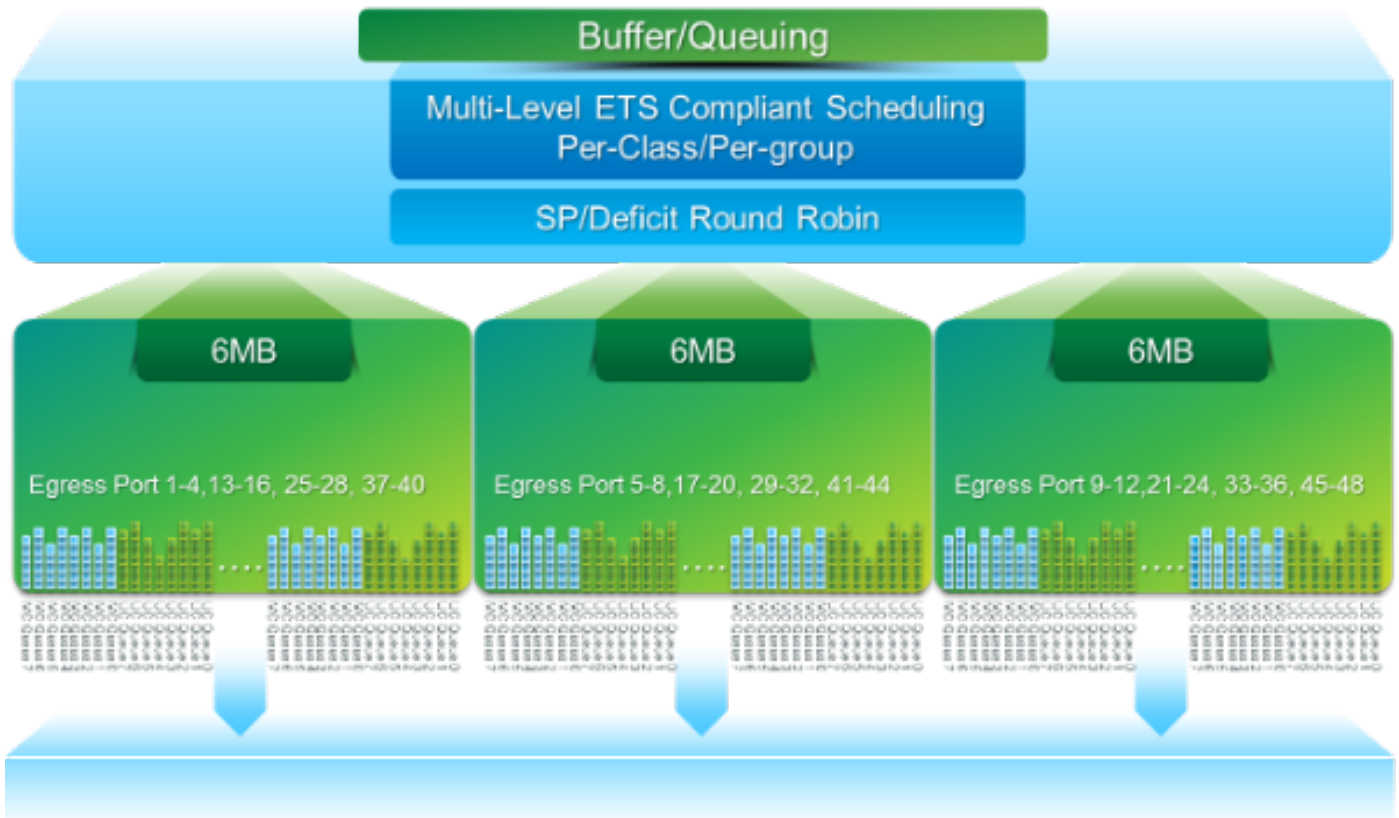
No, esta función no afecta a la latencia ni al rendimiento del dispositivo.

¿Cuál es el impacto del intervalo de sondeo de hardware ABM inferior?

De forma predeterminada, el intervalo de sondeo del hardware es de 4 milisegundos. Puede configurar este valor tan bajo como 10 nanosegundos. No hay impacto en el rendimiento o la latencia debido al intervalo de sondeo de hardware más bajo. Se selecciona el sondeo de hardware predeterminado de 4 milisegundos para asegurarse de que no se desbordan los contadores del histograma antes de que el software sondee cada segundo. Si reduce el intervalo de sondeo de hardware, podría saturar los contadores de hardware a 255 muestras. El dispositivo no puede manejar un sondeo de software inferior a un segundo para hacer coincidir el sondeo de hardware inferior debido a las restricciones de CPU y memoria. El informe técnico tiene el ejemplo del intervalo de sondeo de hardware inferior y su caso práctico.

Apéndice - Información sobre funciones

Administración de búfer

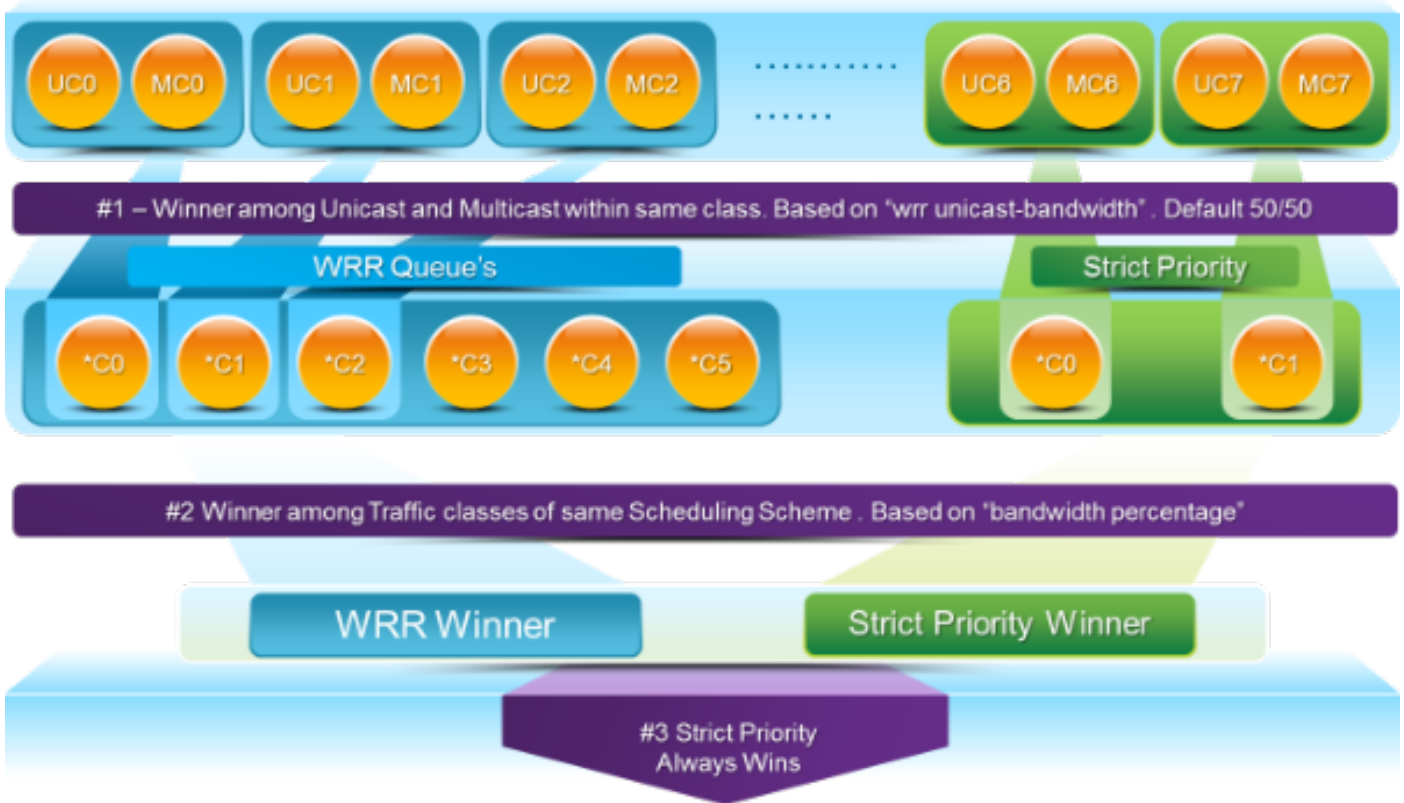


- Búfer de paquetes de 18 MB compartido por tres bloques OB: ~4 MB reservados: Tamaño basado en la unidad máxima de transmisión (MTU) configurada (por suma de puertos de 2 x tamaño de MTU x número de grupos de QoS habilitados)~14 MB compartidos: Resto del búfer total~767 KB de OB: 0 para los paquetes destinados a la CPU
- Un conjunto de 16 puertos comparte 6 MB para cada OB (**comando show hardware internal mtc-usd info port-mapping**)

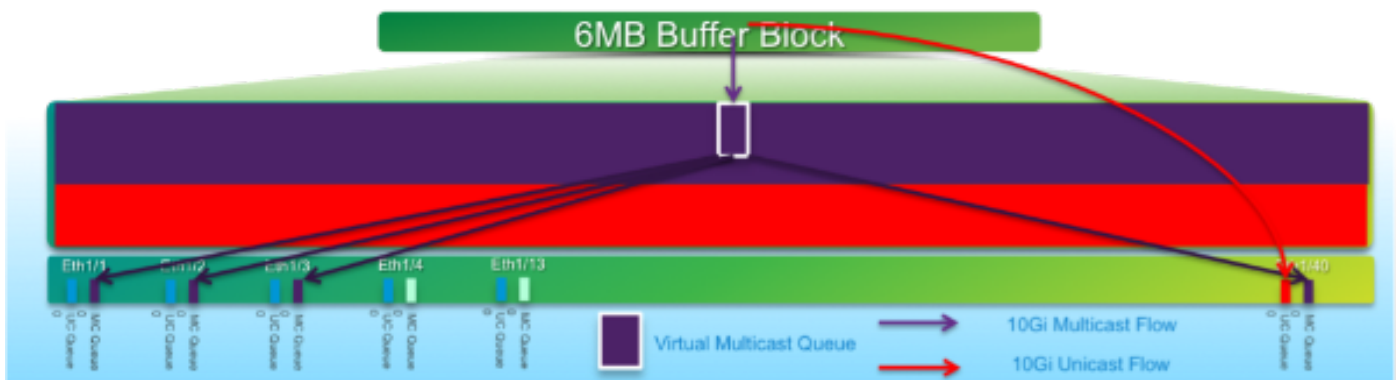
Planificación

Programación de tres capas:

- Unidifusión y multidifusión
- Clases de tráfico del mismo esquema de programación
- Clases de tráfico en todo el esquema



Receptor lento de multidifusión



En este diagrama:

- La congestión sostenida se introduce en 1 G Eth1/40.
- Otros receptores multicast (Eth1/1 - 3) en el bloque buffer se ven afectados debido al comportamiento de programación multicast. *Los receptores de otros bloques de búfer no se ven afectados.*
- El "receptor lento multidifusión" se puede aplicar a e1/40 para evitar la pérdida de tráfico en puertos no congestionados.
- "Receptor lento multidifusión" drena la multidifusión a una velocidad de 10 G en Eth1/40. *Todavía se espera que ocurran caídas en el puerto congestionado.*
- Configurado con el comando **hardware profile multicast slow-receive port <x>**.

Supervisión de búfer activa

Consulte el informe técnico [Supervisión activa de búfer del Cisco Nexus 3548](#) para obtener una descripción general de esta función.

Implementación de hardware

- ASIC tiene 18 cubetas y cada cubeta corresponde a un rango de utilización de búfer (por ejemplo, 0-384 KB, 385-768 KB, etc.).
- ASIC sondea la utilización del búfer para todos los puertos cada 4 milisegundos (valor predeterminado). Este intervalo de sondeo ASIC se puede configurar hasta en 10 nanosegundos.
- Según la utilización del búfer para cada intervalo de sondeo de hardware, se incrementa el contador de depósito para el intervalo correspondiente. Es decir, si el puerto 25 consume 500 KB de búfer, el contador de la cubeta #2 (385-768 KB) se incrementa.
- Este contador de utilización del búfer se mantiene para cada interfaz en formato de histograma.
- Cada cubeta se representa con 8 bits, por lo que el contador se reduce a 255 y se reinicia una vez que el software lee los datos.

Implementación de software

- Cada segundo, el software sondea ASIC para descargar y borrar todos los contadores de histograma.
- Estos contadores de histograma se mantienen en la memoria durante 60 minutos con una granularidad de un segundo.
- El software también se asegura de que copia el histograma del búfer en la memoria flash de inicialización cada hora, que se puede copiar al analizador para un análisis adicional.
- Efectivamente, esto mantiene dos horas de datos del histograma del búfer para todos los puertos, la última hora en la memoria y la segunda hora en la memoria flash de inicialización.