

Nexus 3500 Output Drops e Buffer QoS

Sommario

[Introduzione](#)

[Metodologia](#)

[Controlla rilasci di output](#)

[Determinare se i rilasci sono unicast o multicast](#)

[Determinare il buffer di output utilizzato](#)

[Controllare il monitoraggio attivo del buffer](#)

[Incremento attivo contatori](#)

[Output breve](#)

[Output dettagliato](#)

[Genera un registro al superamento di una soglia](#)

[ID bug Cisco di rilievo](#)

[Domande frequenti](#)

[Appendice - Informazioni sulle feature](#)

[Gestione buffer](#)

[Programmazione](#)

[Multicast Slow-Receiver](#)

[Monitoraggio attivo del buffer](#)

[Implementazione hardware](#)

[Implementazione del software](#)

Introduzione

Questo documento descrive i comandi usati per risolvere i problemi relativi al tipo di traffico scartato sulla piattaforma Nexus 3500 e sul buffer di output (OB) in cui il traffico viene scartato.

Metodologia

1. [Controlla rilasci di output](#)
2. [Determinare se i rilasci sono unicast o multicast](#)
3. [Determinare il buffer di output utilizzato](#)
4. [Controllare il monitoraggio attivo del buffer](#)

Controlla rilasci di output

Controllare le statistiche dell'interfaccia fisica per determinare se il traffico viene scartato nella direzione di uscita. Determinare se il contatore "output scartato" nella direzione TX aumenta e/o è diverso da zero.

```
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

```

Determinare se i rilasci sono unicast o multicast

Dopo aver determinato che l'interfaccia rifiuta il traffico, immettere il comando **show queuing interface <x/y>** per verificare se il traffico viene scartato come multicast o unicast. Nelle versioni precedenti alla 6.0(2)A3(1), l'output è simile al seguente:

```

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

```

Nella release 6.0(2)A3(1) e successive, l'output è simile al seguente:

```
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: Se il ricevitore multicast lento è configurato per la porta, vedere per informazioni sulle funzionalità, le cadute non vengono rilevate con il comando `show queuing interface Eth<x/y>` a causa di un limite hardware. Vedere l'ID bug Cisco [CSCUj21006](https://www.cisco.com/c/enus/bugtools/bugtools/bugtable/CSCUj21006.html).

Determinare il buffer di output utilizzato

In Nexus 3500, ci sono tre pool di buffer utilizzati nella direzione di uscita. L'output del comando `show hardware internal mtc-usd info port-mapping` fornisce le informazioni di mappatura.

```
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 prima parte dei risultati indica che il pool OB 0 viene utilizzato dalle porte anteriori, ad esempio 45, 46, 47, 48 e così via, mentre OB1 viene utilizzato dalle porte anteriori 17, 18 e così via.

La seconda parte dei risultati indica che Eth1/1 è mappato alla porta OB2 12, Eth1/2 alla porta OB2 13 e così via.

La porta in discussione, Eth1/7, è mappata a OB1.

Per ulteriori informazioni, vedere la sezione [Gestione buffer](#) in questo documento.

Controllare il monitoraggio attivo del buffer

Per ulteriori informazioni su questa funzione, vedere il white paper sul [monitoraggio attivo del buffer Cisco Nexus 3548](#) e la sezione riportata in questo documento.


```

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
...

```

Le informazioni di ogni riga vengono registrate a un secondo intervallo. Ogni colonna rappresenta l'utilizzo del buffer. Come accennato nei risultati del comando, se viene segnalato un valore diverso da zero per la colonna "384", significa che l'utilizzo del buffer è stato compreso tra 0 e 384 KByte quando l'ABM ha eseguito il polling dell'utilizzo dell'OB. Il numero diverso da zero è il numero di volte in cui è stato segnalato l'utilizzo.

Questi risultati indicano che OB1 ha utilizzato in media 5,376 MB tra 249 e 253 volte per ogni secondo negli ultimi 10 secondi per Eth1/7. Occorrono 4298 microsecondi (us) per cancellare il buffer di questo traffico.

Genera un registro al superamento di una soglia

Se l'utilizzo del contatore di rilascio e del buffer aumenta periodicamente, è possibile impostare una soglia e generare un messaggio di registro quando la soglia viene superata.

```

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

```

Il comando è impostato per monitorare il traffico unicast a un intervallo di 10 nanosecondi e quando supera il 75% del buffer genera un registro.

È inoltre possibile creare uno scheduler per raccogliere statistiche ABM e output contatore interfaccia ogni ora e aggiungerlo ai file bootflash. Questo esempio è relativo al traffico 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 bug Cisco di rilievo

- ID bug Cisco [CSCum21350](#): I rapidi flap delle porte causano la perdita di tutto il traffico multicast/broadcast TX da parte di tutte le porte nello stesso buffer QoS. Questa condizione viene risolta nella release 6.0(2)A1(1d) e successive.
- ID bug Cisco [CSCuq96923](#): Il blocco del buffer multicast è bloccato, con conseguenti perdite

di multicast/broadcast in uscita. La questione è ancora allo studio.

- ID bug Cisco [CSCva20344](#): Blocco/blocco buffer Nexus 3500 - nessun multicast TX o broadcast. Problema irriproducibile, potenzialmente risolto nelle release 6.0(2)U6(7), 6.0(2)A6(8) e 6.0(2)A8(3).
- ID bug Cisco [CSCvi93997](#): Switch Cisco Nexus 3500 con blocco del buffer di output bloccato. Questo problema è risolto nelle release 7.0(3)I7(8) e 9.3(3).

Domande frequenti

L'ABM influisce sulle prestazioni o sulla latenza?

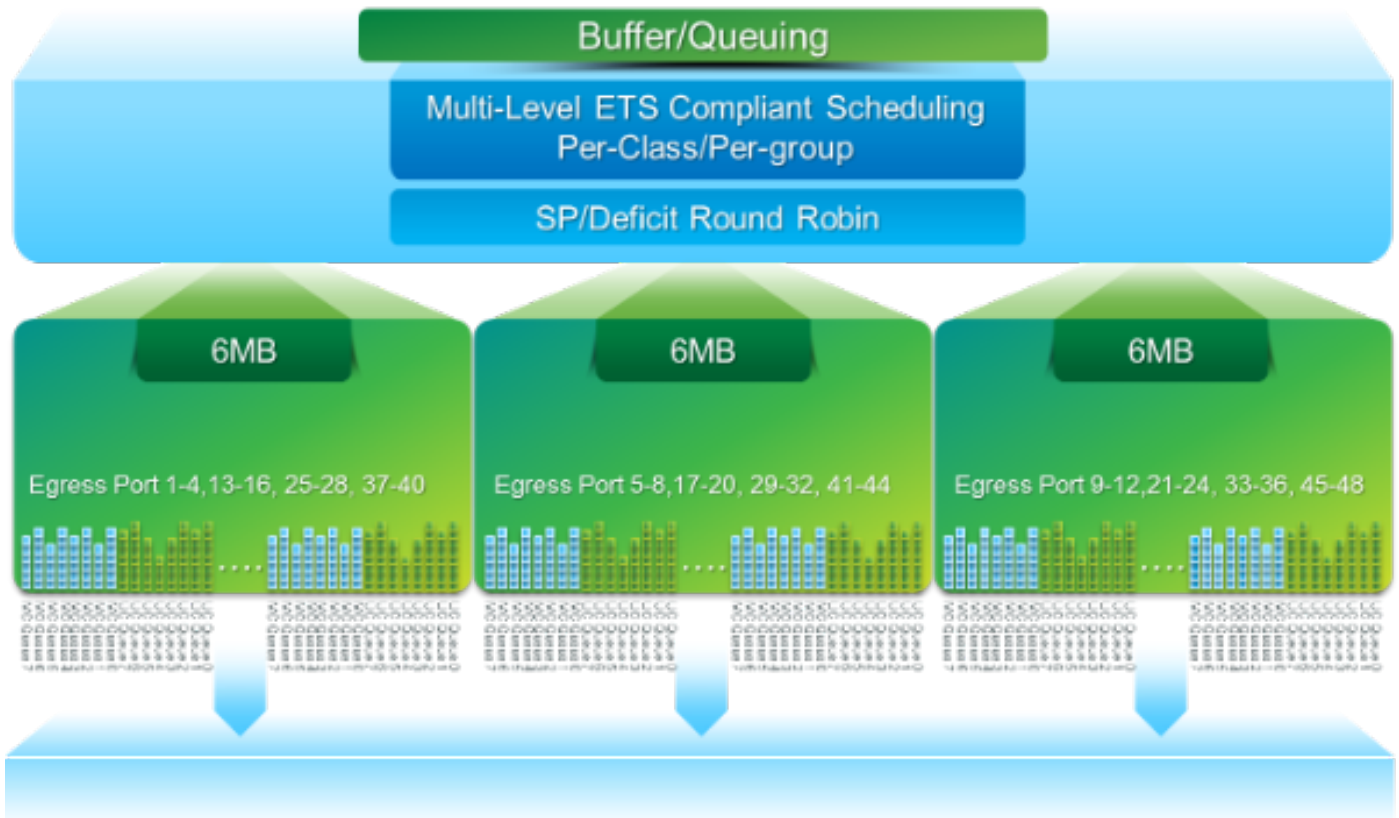
No, questa funzionalità non influisce sulla latenza o sulle prestazioni del dispositivo.

Qual è l'impatto dell'intervallo di polling hardware ABM inferiore?

Per impostazione predefinita, l'intervallo di polling hardware è di 4 millisecondi. È possibile configurare questo valore a 10 nanosecondi. L'intervallo di polling hardware inferiore non ha alcun impatto sulle prestazioni o sulla latenza. È selezionato il polling hardware predefinito di 4 millisecondi per garantire che non si esegua l'overflow dei contatori dell'istogramma prima che il software esegua il polling ogni secondo. Se si riduce l'intervallo di polling hardware, i contatori hardware potrebbero essere saturi a 255 campioni. Il dispositivo non è in grado di gestire un polling software inferiore a un secondo, in modo che corrisponda al polling hardware inferiore a causa di limitazioni della CPU e della memoria. Nel white paper è riportato l'esempio dell'intervallo di polling hardware inferiore e del relativo caso di utilizzo.

Appendice - Informazioni sulle feature

Gestione buffer

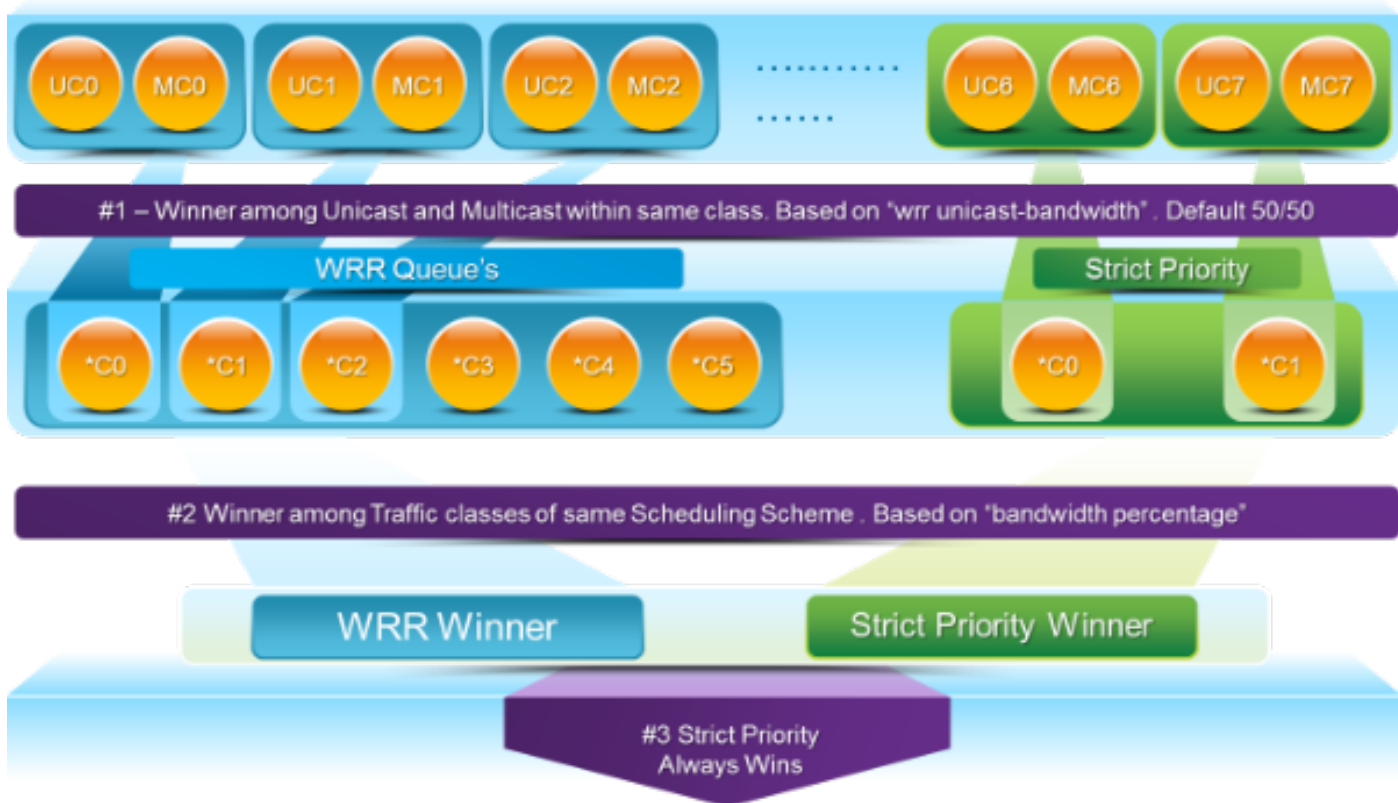


- 18 MB di buffer di pacchetto condiviso da tre blocchi OB: ~4 MB riservati: Dimensione basata sulla MTU (Maximum Transmission Unit) configurata (somma per porta di $2 \times \text{MTU Size} \times \text{n. di gruppi QoS abilitati}$) Circa 14 MB condivisi: Resto del buffer totale Circa 767 KB di OB: 0 per i pacchetti destinati alla CPU
- 6 MB per ciascun OB è condiviso da un set di 16 porte (comando **show hardware internal mtc-usd info port-mapping**)

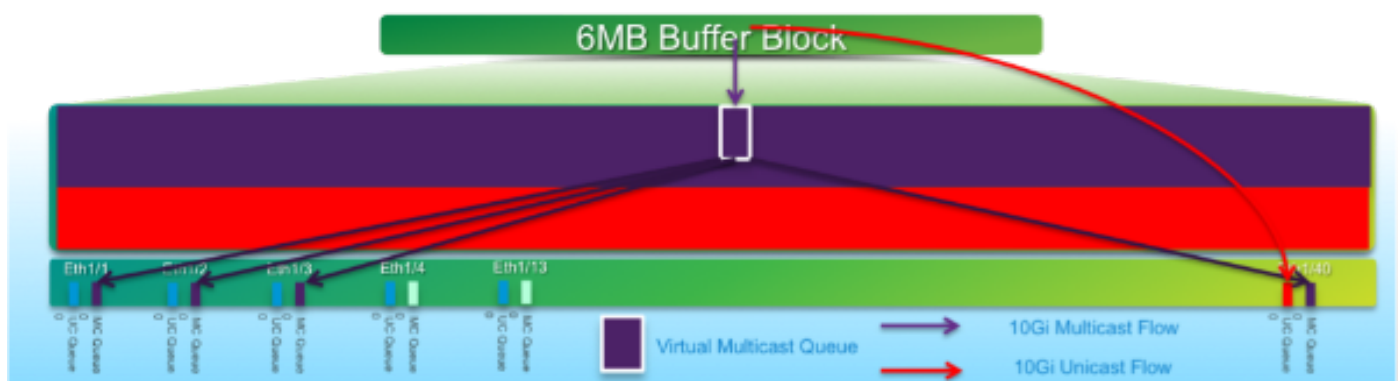
Programmazione

Programmazione a tre livelli:

- Unicast e multicast
- Classi di traffico dello stesso schema di pianificazione
- Classi di traffico nello schema



Multicast Slow-Receiver



In questo diagramma:

- La congestione sostenuta è introdotta su 1 G Eth1/40.
- Altri ricevitori multicast (Eth1/1 - 3) sul blocco buffer sono interessati dal comportamento della programmazione multicast. *I ricevitori di altri blocchi di buffer non vengono modificati.*
- È possibile applicare "Multicast slow-receiver" al modello e1/40 per evitare perdite di traffico sulle porte non congestionate.
- "Multicast slow-receiver" scarica il multicast a una velocità di 10 G su Eth1/40. *Ci si aspetta ancora che si verifichino delle cadute sulla porta congestionata.*
- Configurato con il comando **hardware profile multicast slow-receiver port <x>**.

Monitoraggio attivo del buffer

Per una panoramica di questa funzione, consultare il white paper sul [monitoraggio attivo del buffer Cisco Nexus 3548](#).

Implementazione hardware

- L'ASIC dispone di 18 bucket e ogni bucket corrisponde a un intervallo di utilizzo del buffer (ad esempio, 0-384 KB, 385-768 KB e così via).
- L'ASIC esegue il polling dell'utilizzo del buffer per tutte le porte ogni 4 millisecondi (impostazione predefinita). Questo intervallo di polling ASIC è configurabile a partire da 10 nanosecondi.
- In base all'utilizzo del buffer per ogni intervallo di polling hardware, il contatore di bucket per l'intervallo corrispondente viene incrementato. Ciò significa che se la porta 25 utilizza 500 KB di buffer, il contatore di bucket n. 2 (385-768 KB) viene incrementato.
- Questo contatore di utilizzo del buffer viene gestito per ogni interfaccia in formato istogramma.
- Ogni bucket è rappresentato da 8 bit, quindi il contatore raggiunge al massimo 255 ed è reimpostato una volta che il software legge i dati.

Implementazione del software

- Ogni secondo, il software effettua il polling dell'ASIC per scaricare e cancellare tutti i contatori dell'istogramma.
- Questi contatori vengono mantenuti in memoria per 60 minuti con una granularità di un secondo.
- Il software si assicura anche di copiare l'istogramma del buffer sul bootflash ogni ora, che può essere copiato sull'analizzatore per ulteriori analisi.
- In pratica, questa operazione consente di conservare due ore di dati dell'istogramma del buffer per tutte le porte, l'ultima ora nella memoria e la seconda ora nel bootflash.