

Informazioni sul syslog e sulla congestione del buffer Nexus 9000 TAHUSD

Sommario

[Introduzione](#)

[Prerequisiti](#)

[Requisiti](#)

[Componenti usati](#)

[Conoscenza dell'architettura di buffer ASIC a scalabilità cloud Cisco Nexus 9000](#)

[Informazioni sulle sottoscrizioni in eccesso e sugli output ignorati](#)

[Informazioni sul syslog BUFFER_THRESHOLD_EXCEEDED](#)

[Informazioni sul contatore dell'interfaccia Output Discards](#)

[Scenario di sovrassegnazione di esempio](#)

[Fasi successive](#)

[Ulteriori informazioni](#)

[Opzioni di configurazione del syslog BUFFER_THRESHOLD_EXCEEDED](#)

[Registri da raccogliere per scenari di congestione della rete](#)

[Monitoraggio dei micro-burst](#)

Introduzione

Questo documento descrive i meccanismi di accodamento e buffering su switch Cisco Nexus serie 9000 dotati di Cisco Cloud Scale ASIC (Application-Specific Integrated Circuit) con software NX-OS. Questo documento descrive anche i sintomi di sovrascrittura delle porte su questa piattaforma, come ad esempio i contatori di interfaccia e i syslog di output diversi da zero che indicano che sono state superate le soglie del buffer.

Prerequisiti

Requisiti

Cisco consiglia di comprendere le nozioni di base della commutazione Ethernet su reti medie condivise e la necessità di eseguire code/buffer in queste reti. Cisco consiglia inoltre di comprendere le nozioni di base di Quality of Service (QoS) e di buffering sugli switch Cisco Nexus. Per ulteriori informazioni, consultare la documentazione qui:

- [Cisco Nexus serie 9000 NX-OS Quality of Service Configuration Guide, versione 10.1\(x\)](#)
- [Cisco Nexus serie 9000 NX-OS Quality of Service Configuration Guide, versione 9.3\(x\)](#)
- [Cisco Nexus serie 9000 NX-OS Quality of Service Configuration Guide, versione 9.2\(x\)](#)
- [Cisco Nexus serie 9000 NX-OS Quality of Service Configuration Guide, versione 7.x](#)

Componenti usati

Per la stesura del documento, sono stati usati switch Cisco Nexus serie 9000 con Cloud Scale ASIC con software NX-OS versione 9.3(8).

La procedura illustrata in questo documento è applicabile solo all'hardware mostrato qui.

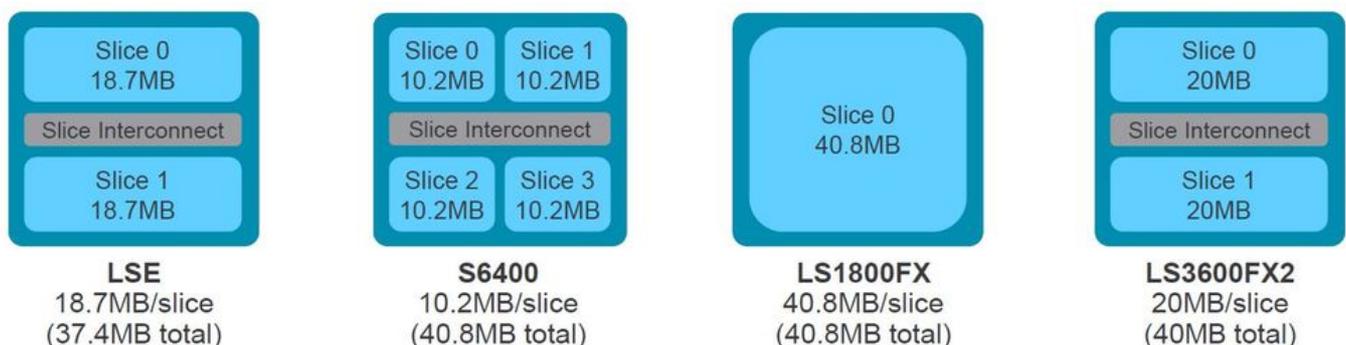
- **Nexus 9200/9300 Fixed Switch** N9K-C92160YC-XN9K-C92300YCN9K-C92304QCN9K-C92348GC-XN9K-C9236CN9K-C9272QN9K-C932CN9K-C9364CN9K-C93108TC-EXN9K-C93108TC-EX-24N9K-C93180LC-EXN9K-C93180YC-EXN9K-C93180YC-EX-24N9K-C93108TC-FXN9K-C93108TC-FX-24N9K-C93180YC-FXN9K-C93180YC-FX-24N9K-C9348GC-FXPN9K-C93240YC-FX2N9K-C93216TC-FX2N9K-C9336C-FX2N9K-C9336C-FX2-EN9K-C93360YC-FX2N9K-C93180YC-FX3N9K-C93108TC-FX3PN9K-C93180YC-FX3SN9K-C9316D-GXN9K-C93600CD-GXN9K-C9364C-GXN9K-C9364D-GX2AN9K-C932D-GX2B
- **Schede di linea per switch modulari Nexus 9500** N9K-X97160YC-EXN9K-X9732C-EXN9K-X9736C-EXN9K-X97284YC-FXN9K-X9732C-FXN9K-X9788TC-FXN9K-X9716D-GX

Le informazioni discusse in questo documento fanno riferimento a dispositivi usati in uno specifico ambiente di emulazione. Su tutti i dispositivi menzionati nel documento la configurazione è stata ripristinata ai valori predefiniti. Se la rete è operativa, valutare attentamente eventuali conseguenze derivanti dall'uso dei comandi.

Conoscenza dell'architettura di buffer ASIC a scalabilità cloud Cisco Nexus 9000

Gli switch Cisco Nexus serie 9000 con Cisco Cloud Scale ASIC implementano un'architettura di buffer in uscita a "memoria condivisa". Un ASIC è diviso in una o più "sezioni". Ogni slice ha un proprio buffer e solo le porte all'interno di quella slice possono usarlo. Fisicamente, ogni sezione è divisa in "celle", che rappresentano porzioni del buffer. Le sezioni vengono suddivise in "gruppi di pool". A ogni gruppo di pool viene allocato un certo numero di celle che non vengono condivise tra gruppi di pool distinti. Ogni gruppo di pool dispone di uno o più "pool", che rappresentano una classe di servizio (CoS) per il traffico unicast o multicast. In questo modo, ciascun gruppo di pool garantisce risorse buffer per i tipi di traffico serviti dal gruppo di pool.

L'immagine mostra visivamente come i vari modelli di Cisco Cloud Scale ASIC sono divisi in sezioni. L'immagine mostra anche come a ciascuna porzione viene allocata una certa quantità di buffer attraverso le celle.

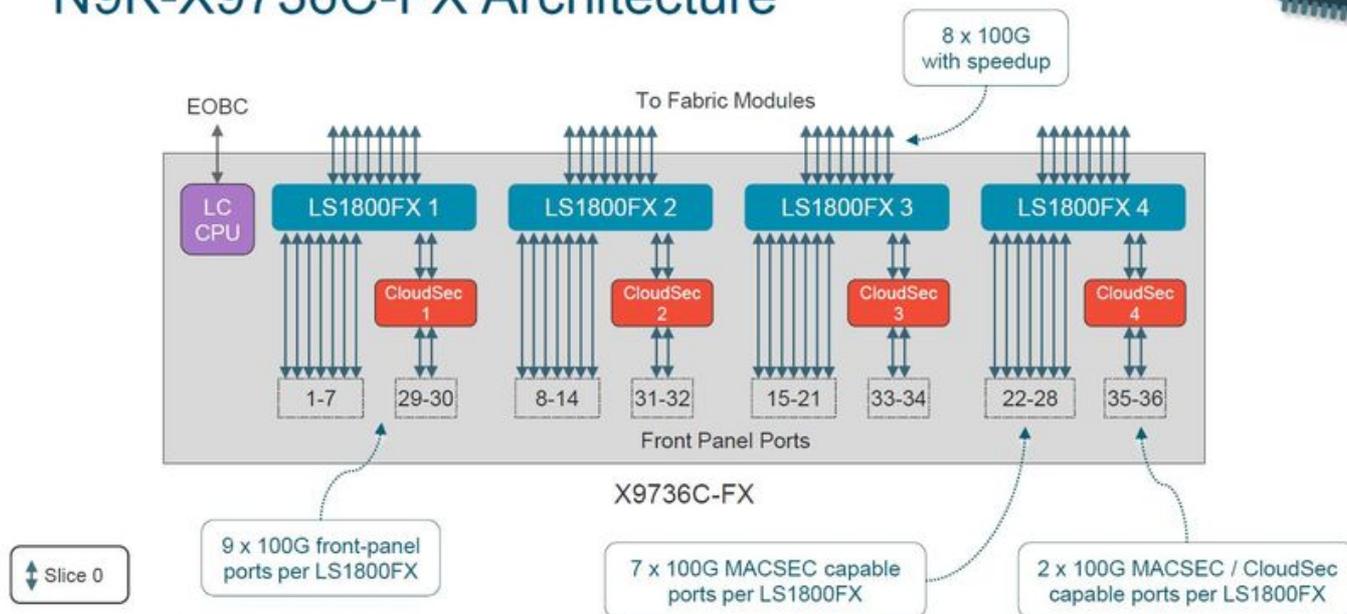


Ogni modello di switch Nexus serie 9000 e di scheda di linea Nexus 9500 ha un numero diverso di ASIC Cisco Cloud Scale all'interno, nonché un layout diverso che determina le porte del pannello anteriore a cui connettersi l'ASIC. Di seguito sono riportati due esempi di utilizzo della scheda di linea N9K-X9736C-FX e dello switch N9K-C9336C-FX2.

La scheda di linea N9K-C9736C-FX dispone di 4 ASIC Cisco Cloud Scale LS1800FX con una slice per ASIC. Internamente, ogni ASIC è definito "unità". A ciascuna slice viene fatto riferimento come "istanza" e viene assegnato un numero intero a base zero che identifica in modo univoco la slice all'interno dello chassis. Il risultato sono le permutazioni mostrate di seguito:

- Unità 0, la slice 0 viene indicata come istanza 0
- Unità 1, la slice 0 viene indicata come istanza 1
- Unità 2: la slice 0 viene indicata come istanza 2
- Unità 3: la slice 0 viene indicata come istanza 3

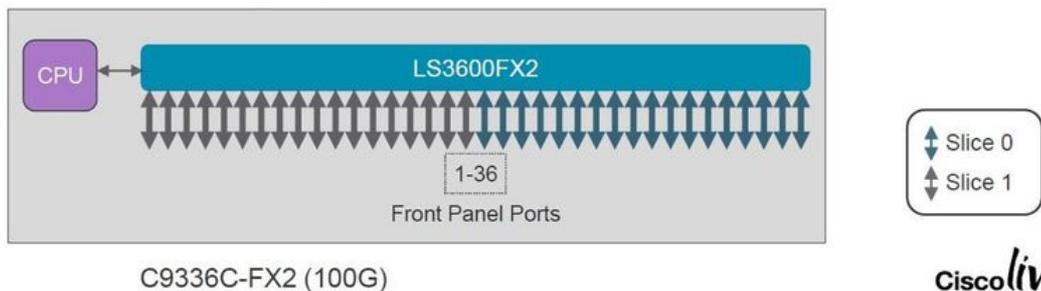
N9K-X9736C-FX Architecture



Lo switch N9K-C936C-FX2 dispone di un ASIC Cisco Cloud Scale LS3600FX2 con due slice per ASIC. Internamente, ogni ASIC è definito "unità". A ciascuna slice viene fatto riferimento come "istanza" e viene assegnato un numero intero a base zero che identifica in modo univoco la slice all'interno dello chassis. Il risultato sono le permutazioni mostrate di seguito:

- Unità 0, la slice 0 viene indicata come istanza 0
- Unità 0, la slice 1 viene indicata come istanza 1

Nexus 9300-FX2 Switch Architecture



Ogni scheda di linea e ogni switch hanno un layout diverso, con numeri di istanza diversi. Comprendere il layout della scheda di linea o dello switch con cui si sta lavorando è importante

per progettare la rete in base a flussi di traffico che richiedono un'elevata larghezza di banda. Il comando **show interface hardware-mappings** può essere usato per correlare ciascuna porta del pannello anteriore a un'unità (ASIC) e a un numero di slice. Di seguito è riportato un esempio di interfaccia Ethernet2/16 di uno switch Nexus 9504 con scheda di linea N9K-X9736C-FX inserita nello slot 2 dello chassis, collegata all'unità 1, slice 0.

```
switch# show interface hardware-mappings
```

Legends:

```

SMod  - Source Mod. 0 is N/A
Unit  - Unit on which port resides. N/A for port channels
HPort - Hardware Port Number or Hardware Trunk Id:
HName - Hardware port name. None means N/A
FPort - Fabric facing port number. 255 means N/A
NPort - Front panel port number
VPort - Virtual Port Number. -1 means N/A
Slice - Slice Number. N/A for BCM systems
SPort - Port Number wrt Slice. N/A for BCM systems
SrcId - Source Id Number. N/A for BCM systems
MacIdx - Mac index. N/A for BCM systems
MacSubPort - Mac sub port. N/A for BCM systems

```

```

-----
-----
Name          Ifindex  Smod  Unit  HPort  FPort  NPort  VPort  Slice  SPort  SrcId  MacId  MacSP  VIF  Block
BlkSrcID
-----
-----
Eth2/1       1a080000  5    0    16    255    0     -1    0     16    32    4     0     145  0
32
Eth2/2       1a080200  5    0    12    255    4     -1    0     12    24    3     0     149  0
24
Eth2/3       1a080400  5    0    8     255    8     -1    0     8     16    2     0     153  0
16
Eth2/4       1a080600  5    0    4     255    12    -1    0     4     8     1     0     157  0     8
Eth2/5       1a080800  5    0    0     255    16    -1    0     0     0     0     0     161  0     0
Eth2/6       1a080a00  5    0    56    255    20    -1    0     56    112   14    0     165  1
40
Eth2/7       1a080c00  5    0    52    255    24    -1    0     52    104   13    0     169  1
32
Eth2/8       1a080e00  6    1    16    255    28    -1    0     16    32    4     0     173  0
32
Eth2/9       1a081000  6    1    12    255    32    -1    0     12    24    3     0     177  0
24
Eth2/10      1a081200  6    1    8     255    36    -1    0     8     16    2     0     181  0
16
Eth2/11      1a081400  6    1    4     255    40    -1    0     4     8     1     0     185  0     8
Eth2/12      1a081600  6    1    0     255    44    -1    0     0     0     0     0     189  0     0
Eth2/13      1a081800  6    1    56    255    48    -1    0     56    112   14    0     193  1
40
Eth2/14      1a081a00  6    1    52    255    52    -1    0     52    104   13    0     197  1
32
Eth2/15      1a081c00  7    2    16    255    56    -1    0     16    32    4     0     201  0
32
Eth2/16    1a081e00  7    2   12    255    60    -1    0   12    24    3     0     205  0
24
Eth2/17      1a082000  7    2    8     255    64    -1    0     8     16    2     0     209  0
16
Eth2/18      1a082200  7    2    4     255    68    -1    0     4     8     1     0     213  0     8
Eth2/19      1a082400  7    2    0     255    72    -1    0     0     0     0     0     217  0     0
Eth2/20      1a082600  7    2    56    255    76    -1    0     56    112   14    0     221  1
40

```

Eth2/21 32	1a082800	7	2	52	255	80	-1	0	52	104	13	0	225	1	
Eth2/22 32	1a082a00	8	3	16	255	84	-1	0	16	32	4	0	229	0	
Eth2/23 24	1a082c00	8	3	12	255	88	-1	0	12	24	3	0	233	0	
Eth2/24 16	1a082e00	8	3	8	255	92	-1	0	8	16	2	0	237	0	
Eth2/25	1a083000	8	3	4	255	96	-1	0	4	8	1	0	241	0	8
Eth2/26	1a083200	8	3	0	255	100	-1	0	0	0	0	0	245	0	0
Eth2/27 40	1a083400	8	3	56	255	104	-1	0	56	112	14	0	249	1	
Eth2/28 32	1a083600	8	3	52	255	108	-1	0	52	104	13	0	253	1	
Eth2/29 24	1a083800	5	0	48	255	112	-1	0	48	96	12	0	257	1	
Eth2/30 16	1a083a00	5	0	44	255	116	-1	0	44	88	11	0	261	1	
Eth2/31 24	1a083c00	6	1	48	255	120	-1	0	48	96	12	0	265	1	
Eth2/32 16	1a083e00	6	1	44	255	124	-1	0	44	88	11	0	269	1	
Eth2/33 24	1a084000	7	2	48	255	128	-1	0	48	96	12	0	273	1	
Eth2/34 16	1a084200	7	2	44	255	132	-1	0	44	88	11	0	277	1	
Eth2/35 24	1a084400	8	3	48	255	136	-1	0	48	96	12	0	281	1	
Eth2/36 16	1a084600	8	3	44	255	140	-1	0	44	88	11	0	285	1	

Informazioni sulle sottoscrizioni in eccesso e sugli output ignorati

Le interfacce collegate a una rete Ethernet possono trasmettere un solo pacchetto alla volta. Quando due pacchetti devono uscire contemporaneamente da un'interfaccia Ethernet, l'interfaccia Ethernet inizia a trasmettere un pacchetto mentre memorizza l'altro. Al termine della trasmissione del primo pacchetto, l'interfaccia Ethernet inizia a trasmettere il secondo pacchetto dal buffer. Quando la somma totale del traffico che deve uscire da un'interfaccia supera la larghezza di banda dell'interfaccia, l'interfaccia viene considerata *sovrascritta*. Ad esempio, se un totale di 15 Gb/s di traffico entra istantaneamente nello switch e deve uscire da un'interfaccia a 10 Gb/s, l'interfaccia a 10 Gb/s ha una sottoscrizione eccessiva in quanto non è in grado di trasmettere 15 Gb/s di traffico alla volta.

Uno switch Cisco Nexus serie 9000 con ASIC a scala cloud gestisce questa contesa tra risorse memorizzando il traffico all'interno dei buffer della slice ASIC associata all'interfaccia di uscita. Se la somma totale del traffico in uscita da un'interfaccia supera la larghezza di banda dell'interfaccia per un periodo di tempo esteso, i buffer della slice ASIC iniziano a riempire di pacchetti che devono uscire dall'interfaccia.

Quando i buffer della slice ASIC raggiungono il 90% di utilizzo, lo switch genera un syslog simile a quello mostrato di seguito:

```
%TAHUSD-SLOT2-4-BUFFER_THRESHOLD_EXCEEDED: Module 2 Instance 0 Pool-group buffer 90 percent threshold is exceeded!
```

Quando i buffer della slice ASIC sono completamente pieni, lo switch scarta il traffico aggiuntivo

che deve uscire dall'interfaccia finché lo spazio nei buffer non diventa libero. Quando lo switch scarta il traffico, lo switch incrementa il contatore Output Discard sull'interfaccia di uscita.

Il contatore syslog generato e gli scarti di output diversi da zero sono entrambi sintomi di un'interfaccia sovrascritta. Ciascun sintomo viene analizzato più dettagliatamente nelle sottosezioni qui.

Informazioni sul syslog BUFFER_THRESHOLD_EXCEEDED

Di seguito è riportato un esempio del syslog BUFFER_THRESHOLD_EXCEEDED.

```
%TAHUSD-SLOTX-4-BUFFER_THRESHOLD_EXCEEDED: Module X Instance Y Pool-group buffer Z percent threshold is exceeded!
```

Questo syslog contiene tre informazioni chiave:

1. **Modulo X** - Slot della scheda di linea contenente l'interfaccia sovrascritta.
2. **Istanza Y** - Numero di istanza assegnato all'ASIC e alla tupla di sezione contenente l'interfaccia sovrascritta.
3. **Buffer Z pool-group**: la soglia del buffer del gruppo pool interessato prima che venga generato il syslog. Si tratta di una percentuale derivata dalle celle usate divisa per le celle totali, come osservato nell'output di **show hardware buffer info pkt-stats** quando collegato al modulo X.

Informazioni sul contatore dell'interfaccia Output Discards

Il contatore dell'interfaccia Output Discards indica il numero di pacchetti ignorati che *avrebbero dovuto* uscire dall'interfaccia, ma che non sono stati in grado di farlo perché il buffer della slice ASIC è pieno e non è in grado di accettare nuovi pacchetti. Il contatore Output Discards è visibile nell'output del comando **show interface** e **mostra gli errori dei contatori dell'interfaccia**, come mostrato di seguito.

```
switch# show interface Ethernet1/1
Ethernet1/1 is up
admin state is up, Dedicated Interface
  Hardware: 1000/10000/25000/40000/50000/100000 Ethernet, address: 7cad.4f6d.f6d8 (bia
7cad.4f6d.f6d8)
  MTU 1500 bytes, BW 40000000 Kbit , DLY 10 usec
  reliability 255/255, txload 232/255, rxload 1/255
  Encapsulation ARPA, medium is broadcast
  Port mode is trunk
  full-duplex, 40 Gb/s, media type is 40G
  Beacon is turned off
  Auto-Negotiation is turned on FEC mode is Auto
  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 off
  Last link flapped 03:16:50
  Last clearing of "show interface" counters never
  3 interface resets
```

```

Load-Interval #1: 30 seconds
 30 seconds input rate 0 bits/sec, 0 packets/sec
 30 seconds output rate 36503585488 bits/sec, 3033870 packets/sec
 input rate 0 bps, 0 pps; output rate 36.50 Gbps, 3.03 Mpps
Load-Interval #2: 5 minute (300 seconds)
 300 seconds input rate 32 bits/sec, 0 packets/sec
 300 seconds output rate 39094683384 bits/sec, 3249159 packets/sec
 input rate 32 bps, 0 pps; output rate 39.09 Gbps, 3.25 Mpps
RX
 0 unicast packets 208 multicast packets 9 broadcast packets
217 input packets 50912 bytes
 0 jumbo packets 0 storm suppression bytes
 0 runs 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
38298127762 unicast packets 6118 multicast packets 0 broadcast packets
38298133880 output packets 57600384931480 bytes
 0 jumbo packets
 0 output error 0 collision 0 deferred 0 late collision
 0 lost carrier 0 no carrier 0 babble 57443534227 output discard <<< Output discards
due to oversubscription
 0 Tx pause

```

```
switch# show interface Ethernet1/1 counters errors
```

```

-----
Port          Align-Err    FCS-Err    Xmit-Err    Rcv-Err    UnderSize  OutDiscards
-----
Eth1/1                0          0          0          0          0  57443534227
-----

Port          Single-Col  Multi-Col   Late-Col   Exces-Col   Carri-Sen   Runts
-----
Eth1/1                0          0          0          0          0          0
-----

Port          Giants SQETest-Err Deferred-Tx IntMacTx-Er IntMacRx-Er Symbol-Err
-----
Eth1/1                0          --          0          0          0          0
-----

Port          InDiscards
-----
Eth1/1                0

```

Scenario di sovrassegnazione di esempio

Si consideri uno scenario in cui il traffico tra due generatori di traffico IXIA attraversa uno switch Nexus 9504 con due schede di linea N9K-X9736C-FX inserite negli slot 1 e 2 dello chassis. 100 Gb/s di traffico entra nello switch tramite l'interfaccia 100 Gb/s Ethernet1/1 e deve uscire dall'interfaccia 40 Gb/s Ethernet2/2. Pertanto, l'iscrizione a Ethernet2/2 è eccessiva. Di seguito è riportata una topologia dello scenario.

93554 93554 |

Occupancy drops		51152555398	0	0	0	0	0
0	0						
AQM drops		0	0	0	0	0	0
0	0						

 -----|

		Output MC Pool counters					
		Pool 0	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
Pool 6	Pool 7						

Dynamic Threshold (cells)		93554	93554	93554	93554	93554	93554
93554	93554						
Dynamic Threshold (desc)		93554	93554	93554	93554	93554	93554
93554	93554						
Dynamic Threshold (inq thr)		64035	64035	64035	64035	64035	64035
64035	64035						

Occupancy drops		0	0	0	0	0	0
0	0						

-----+-----+-----+-----+-----+-----+-----+-----+
 |
 | Additional counters
 |-----+-----+-----+-----+-----+-----+-----+-----+
 |

MEM cell drop reason	:	0
MEM descriptor drop reason	:	0
OPG cell drop reason	:	0
OPG descriptor drop reason	:	0
OPG CPU cell drop reason	:	0
OPG CPU descriptor drop reason	:	0
OPG SPAN cell drop reason	:	0
OPG SPAN descriptor drop reason	:	0
OPOOL cell drop reason	:	0
OPOOL descriptor drop reason	:	0
UC OQUEUE cell drop reason	:	51152556479
MC OQUEUE cell drop reason	:	27573307
OQUEUE descriptor drop reason	:	0
MC OPOOL cell drop reason	:	0
FWD DROP	:	15
SOD	:	0
BMM BP	:	0
No Drop	:	0
Packets received	:	87480806439
TRUNC MTU	:	0
TRUNK BMM BP	:	0
VOQFC messages sent	:	0
SOD messages sent	:	0
SPAN descptor drop	:	0

Unit: 1 Slice: 0

=====

 -----|

		Output Pool-Group drops				
		Drop-PG	No-drop	CPU--PG	LCPU-PG	RCPU-PG
SPAN-PG						

Occupancy drops		0	0	0	0	0
-----------------	--	---	---	---	---	---

0						
AQM drops		0	0	N/A	N/A	N/A
N/A						

```

-----
-----|
|
|                                     Output UC Pool counters
|
|                                     Pool 0   Pool 1   Pool 2   Pool 3   Pool 4   Pool 5
Pool 6   Pool 7 |
-----
-----|
Dynamic Threshold (cells)           93554   93554   93554   93554   93554   93554
93554   93554 |
Occupancy drops                     0       0       0       0       0       0
0       0 |
AQM drops                           0       0       0       0       0       0
0       0 |

```

```

-----
-----|
|
|                                     Output MC Pool counters
|
|                                     Pool 0   Pool 1   Pool 2   Pool 3   Pool 4   Pool 5
Pool 6   Pool 7 |
-----
-----|
Dynamic Threshold (cells)           93554   93554   93554   93554   93554   93554
93554   93554 |
Dynamic Threshold (desc)           93554   93554   93554   93554   93554   93554
93554   93554 |
Dynamic Threshold (inq thr)        64035   64035   64035   64035   64035   64035
64035   64035 |
Occupancy drops                     0       0       0       0       0       0
0       0 |

```

```

-----+-----+-----+-----+-----+-----+-----+-----+
|
|                                     Additional counters
|
-----+-----+-----+-----+-----+-----+-----+
MEM cell drop reason                :      0
MEM descriptor drop reason           :      0
OPG cell drop reason                 :      0
OPG descriptor drop reason           :      0
OPG CPU cell drop reason              :      0
OPG CPU descriptor drop reason       :      0
OPG SPAN cell drop reason            :      0
OPG SPAN descriptor drop reason      :      0
OPOOL cell drop reason               :      0
OPOOL descriptor drop reason         :      0
UC OQUEUE cell drop reason           :      0
MC OQUEUE cell drop reason           :      0
OQUEUE descriptor drop reason        :      0
MC OPOOL cell drop reason            :      0
FWD DROP                             :      8
SOD                                  :      0
BMM BP                               :      0
No Drop                              :      0
Packets received                     : 45981341
TRUNC MTU                            :      0
TRUNK BMM BP                         :      0
VOQFC messages sent                  :      0
SOD messages sent                    :      0

```

Tenere presente che ogni tupla di unità/porzione ASIC è rappresentata da un'unica istanza identificata. L'output del comando **show hardware internal buffer info pkt-stats** visualizza informazioni dettagliate sul gruppo di pool congestionato (abbreviato come "PG") per ciascuna istanza. Il comando mostra anche il numero massimo/massimo storico di celle nel buffer che sono state utilizzate. Infine, il comando mostra uno snapshot istantaneo degli identificatori di porta ASIC di scala cloud delle porte con traffico memorizzato nel buffer. Di seguito è riportato un esempio di questo comando.

```
switch# attach module 2
module-2# show hardware internal buffer info pkt-stats
```

Instance 0

```
=====
|-----|
|-----|
|                                     Output Pool-Group Buffer Utilization (cells/desc)
|
|                                     Drop-PG      No-drop      CPU--PG      LCPU-PG      RCPU-PG
SPAN-PG |-----|
|-----|
|
Total Instant Usage (cells)      59992          0          0          0          0
0 |
Remaining Instant Usage (cells)   33562          0          1500       250         1500
1500 |
Peak/Max Cells Used               90415          0          N/A         N/A         N/A
N/A |
Switch Cells Count               93554          0          1500       250         1500
1500 |

Total Instant Usage (desc)          0          0          0          0          0
0 |
Remaining Instant Usage (desc)      93554       0          1500       250         1500
1500 |
Switch Desc Count                   93554       0          1500       250         1500
1500 |
```

```
|-----|
|-----|
|                                     Output UC Pool Buffer Utilization (cells/desc)
|
|                                     Pool 0      Pool 1      Pool 2      Pool 3      Pool 4      Pool 5
Pool 6      Pool 7 |-----|
|-----|
|
Total Instant Usage (cells)      60027          0          0          0          0          0
0 |
Total Instant Usage (desc)          0          0          0          0          0          0
0 |
Peak/Max Cells Used               62047          0          0          0          0          0
0 |
```

```
|-----|
|-----|
|                                     Output MC Pool Buffer Utilization (cells/desc)
|
|                                     Pool 0      Pool 1      Pool 2      Pool 3      Pool 4      Pool 5
```

```

Pool 6      Pool 7 |
|-----|
|-----|
Total Instant Usage (cells)      0      0      0      0      0      0
0      0 |
Total Instant Usage (desc)      0      0      0      0      0      0
0      0 |
Total Instant Usage (inq cells)  0      0      0      0      0      0
0      0 |
Total Instant Usage (packets)    0      0      0      0      0      0
0      0 |
Peak/Max Cells Used              60399    0      0      0      0      0
0      0 |

```

```

|-----|
| Instant Buffer utilization per queue per port |
| Each line displays the number of cells/desc utilized for a given |
| port for each QoS queue |
| One cell represents approximately 416 bytes |
|-----|

```

```

|-----+-----+-----+-----+-----+-----+-----+-----+
|ASIC Port      Q7      Q6      Q5      Q4      Q3      Q2      Q1      Q0      |
|-----+-----+-----+-----+-----+-----+-----+-----+

```

```

[12]
<<< ASIC Port 12 in Unit 0 Instance 0 is likely the congested egress interface
      UC->      0      0      0      0      0      0      0      59988
MC cells->     0      0      0      0      0      0      0      0
MC desc->      0      0      0      0      0      0      0      0

```

Vedere anche la variazione del picco del comando. Utilizzare questo comando per associare il syslog a un potenziale picco in un particolare gruppo di pool, pool o porta

```
switch# show hardware internal buffer info pkt-stats peak
```

```
slot 1
=====
```

```
Instance 0
```

```
=====
```

```

|-----+-----+-----+-----+-----+-----+
| Pool-Group Peak counters |
|-----+-----+-----+-----+-----+-----+
Drop PG      :      0
No-drop PG   :      0

```

```

|-----+-----+-----+-----+-----+-----+
| Pool Peak counters |
|-----+-----+-----+-----+-----+-----+
MC Pool 0    :      0
MC Pool 1    :      0
MC Pool 2    :      0
MC Pool 3    :      0
MC Pool 4    :      0
MC Pool 5    :      0
MC Pool 6    :      0
MC Pool 7    :      0

```

```

UC Pool 0      :      0
UC Pool 1      :      0
UC Pool 2      :      0
UC Pool 3      :      0
UC Pool 4      :      0
UC Pool 5      :      0
UC Pool 6      :      0
UC Pool 7      :      0

```

```

|-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Port Peak counters                             |
|      classes mapped to count_0: 0 1 2 3 4 5 6 7                               |
|      classes mapped to count_1: None                                           |
|-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

```

[0]                                     <<< ASIC Port. This can be checked via "show
hardware interface-mappings"
      count_0      :      0
      count_1      :      0
[1]
      count_0      :      0
      count_1      :      0

```

Il comando **show interface hardware-mapping** può essere usato per convertire l'identificatore della porta ASIC su scala cloud in una porta del pannello anteriore. Nell'esempio sopra riportato, la porta ASIC 12 (rappresentata dalla colonna HPort nell'output del comando **show interface hardware-mapping**) associata all'unità ASIC 0 sulla slice/istanza 0 contiene 59.988 celle occupate di 416 byte ciascuna. Di seguito è riportato un esempio di comando **show interface hardware-mappings** che permette di mappare l'interfaccia alla porta Ethernet2/2 del pannello anteriore.

```

switch# show interface hardware-mappings
Legends:

```

- SMod - Source Mod. 0 is N/A
- Unit - Unit on which port resides. N/A for port channels
- HPort - Hardware Port Number or Hardware Trunk Id:
- HName - Hardware port name. None means N/A
- FPort - Fabric facing port number. 255 means N/A
- NPort - Front panel port number
- VPort - Virtual Port Number. -1 means N/A
- Slice - Slice Number. N/A for BCM systems
- SPort - Port Number wrt Slice. N/A for BCM systems
- SrcId - Source Id Number. N/A for BCM systems
- MacIdx - Mac index. N/A for BCM systems
- MacSubPort - Mac sub port. N/A for BCM systems

```

-----
-----
Name      Ifindex  Smod Unit HPort FPort NPort VPort Slice SPort SrcId MacId MacSP VIF  Block
BlkSrcID
-----
-----
Eth2/2    1a080200 5    0    12   255  4    -1    0    12   24   3    0    149  0
24

```

È possibile correlare ulteriormente la sovrascrittura dell'interfaccia Ethernet2/2 con le perdite di coda QoS con il comando **show queuing interface**. Di seguito è riportato un esempio.

```

switch# show queuing interface Ethernet2/2

```

Egress Queuing for Ethernet2/2 [System]

QoS-Group#	Bandwidth%	PrioLevel	Min	Shape Max	Units	QLimit
7	-	1	-	-	-	9 (D)
6	0	-	-	-	-	9 (D)
5	0	-	-	-	-	9 (D)
4	0	-	-	-	-	9 (D)
3	0	-	-	-	-	9 (D)
2	0	-	-	-	-	9 (D)
1	0	-	-	-	-	9 (D)
0	100	-	-	-	-	9 (D)

+-----+
 | QOS GROUP 0 |

	Unicast	Multicast
Tx Pkts	35593332351	18407162
Tx Byts	53532371857088	27684371648
WRED/AFD & Tail Drop Pkts	53390604466	27573307
WRED/AFD & Tail Drop Byts	80299469116864	110293228
Q Depth Byts	24961664	0
WD & Tail Drop Pkts	53390604466	27573307

+-----+
 | QOS GROUP 1 |

	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0

+-----+
 | QOS GROUP 2 |

	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0

+-----+
 | QOS GROUP 3 |

	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0

+-----+
 | QOS GROUP 4 |

	Unicast	Multicast

Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0

QOS GROUP 5

	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0

QOS GROUP 6

	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0

QOS GROUP 7

	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0

CONTROL QOS GROUP

	Unicast	Multicast
Tx Pkts	5704	0
Tx Byts	725030	0
Tail Drop Pkts	0	0
Tail Drop Byts	0	0

SPAN QOS GROUP

	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0

Per Slice Egress SPAN Statistics

SPAN Copies Tail Drop Pkts	0
SPAN Input Queue Drop Pkts	0
SPAN Copies/Transit Tail Drop Pkts	0
SPAN Input Desc Drop Pkts	0

Infine, è possibile verificare che l'interfaccia Ethernet2/2 in uscita abbia un contatore dei rifiuti di output diverso da zero con il comando **show interface**. Di seguito è riportato un esempio.

```
switch# show interface Ethernet2/2
Ethernet2/2 is up
admin state is up, Dedicated Interface
  Hardware: 1000/10000/25000/40000/50000/100000 Ethernet, address: 7cad.4f6d.f6d8 (bia
7cad.4f6d.f6d8)
  MTU 1500 bytes, BW 40000000 Kbit , DLY 10 usec
  reliability 255/255, txload 232/255, rxload 1/255
  Encapsulation ARPA, medium is broadcast
  Port mode is trunk
  full-duplex, 40 Gb/s, media type is 40G
  Beacon is turned off
  Auto-Negotiation is turned on FEC mode is Auto
  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 off
  Last link flapped 03:16:50
  Last clearing of "show interface" counters never
  3 interface resets
  Load-Interval #1: 30 seconds
    30 seconds input rate 0 bits/sec, 0 packets/sec
    30 seconds output rate 36503585488 bits/sec, 3033870 packets/sec
    input rate 0 bps, 0 pps; output rate 36.50 Gbps, 3.03 Mpps
  Load-Interval #2: 5 minute (300 seconds)
    300 seconds input rate 32 bits/sec, 0 packets/sec
    300 seconds output rate 39094683384 bits/sec, 3249159 packets/sec
    input rate 32 bps, 0 pps; output rate 39.09 Gbps, 3.25 Mpps
RX
  0 unicast packets  208 multicast packets  9 broadcast packets
  217 input packets  50912 bytes
  0 jumbo packets  0 storm suppression bytes
  0 runs  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
  38298127762 unicast packets  6118 multicast packets  0 broadcast packets
  38298133880 output packets  57600384931480 bytes
  0 jumbo packets
  0 output error  0 collision  0 deferred  0 late collision
  0 lost carrier  0 no carrier  0 babble  57443534227 output discard      <<< Output discards
due to oversubscription
  0 Tx pause
```

Fasi successive

Se si osservano scarti di output su uno switch Nexus serie 9000 con ASIC Cloud Scale, è possibile risolvere il problema utilizzando uno o più dei seguenti metodi:

- Se l'interfaccia di cui si verificano gli scarti di output è un'unica interfaccia e non è inclusa in

un canale porta, l'aggiornamento della larghezza di banda dell'interfaccia può aiutare a ridurre la congestione. Ad esempio, se un'interfaccia in uscita congestionata è un'interfaccia a 10 Gbps, l'aggiornamento a un'interfaccia a 25 Gbps, 40 Gbps o 100 Gbps può aiutare a risolvere il problema. A seconda del fattore di forma del ricetrasmittitore dell'interfaccia di uscita, questo può essere fatto aggiornando il ricetrasmittitore (ad esempio, migrando da un SFP+ 10 Gbps inserito in un CVR-QSFP-SFP10G all'interno di una porta QSFP a un ricetrasmittitore QSFP nativo 40 Gbps). Questa operazione può essere eseguita anche migrando la configurazione dell'interfaccia di uscita congestionata da una porta a 10 Gbps a una porta a 25 Gbps, 40 Gbps o 100 Gbps.

- Se l'interfaccia di cui si verificano gli scarti di output è un'unica interfaccia e non è inclusa in un canale porta, configurare l'interfaccia congestionata in modo che sia membro di un canale porta con un'altra interfaccia della stessa larghezza di banda può ridurre la congestione.
- Se l'interfaccia di cui si verificano gli output ignorati è un'interfaccia porta-canale, l'aggiunta di ulteriori membri al canale porta può aumentare la larghezza di banda del canale porta complessivo e migliorare l'hashing con bilanciamento del carico per più flussi di traffico di grandi dimensioni.
- Verificare se i flussi di traffico congestionato tra gli host della rete interessano interfacce con velocità inferiore (ad esempio, il traffico in entrata su uno switch tramite un'interfaccia a 40 Gb/s ed in uscita su uno switch tramite un'interfaccia a 10 Gb/s). Questo può essere un collo di bottiglia che causa congestione della rete. Eliminare questo collo di bottiglia aggiornando l'interfaccia a velocità inferiore (ad esempio, 10 Gbps) a un'interfaccia a velocità superiore (ad esempio, 25 Gbps, 40 Gbps e così via) può ridurre la congestione della rete.
- Se non è possibile aumentare la larghezza di banda disponibile sull'interfaccia di uscita congestionata, convalidare il [QoS end-to-end](#) e applicare le azioni di accodamento appropriate per la rete.
- Se i micro-burst sono una potenziale causa di congestione intermittente, fare riferimento alla sezione [Monitoraggio dei micro-burst](#) di questo documento per informazioni sulla configurazione del monitoraggio dei micro-burst.

Ulteriori informazioni

In questa sezione del documento vengono fornite informazioni aggiuntive sui passaggi successivi da eseguire quando si incontra il syslog BUFFER_THRESHOLD_EXCEEDED, gli scenari di congestione/sovrascrittura della rete e i contatori incrementali dell'interfaccia di eliminazione dell'output.

Opzioni di configurazione del syslog BUFFER_THRESHOLD_EXCEEDED

È possibile modificare l'intervallo di polling dello stato del buffer di sistema, che controlla la frequenza con cui il sistema esegue il polling dell'utilizzo corrente dei buffer delle slice ASIC. A tale scopo, usare il comando **hardware profile buffer info poll-interval** in modalità di configurazione globale. Il valore di configurazione predefinito è 5.000 millisecondi. Questa configurazione può essere modificata globalmente o per modulo. Di seguito è riportato un esempio di questo comando di configurazione, con il quale è stato modificato in un valore di 1.000 millisecondi.

```
switch# configure terminal  
Enter configuration commands, one per line. End with CNTL/Z.  
switch(config)# hardware profile buffer info poll-interval timer 1000
```

```
switch(config)# end
switch# show running-config | include hardware.profile.buffer
hardware profile buffer info poll-interval timer 1000
switch#
```

È possibile modificare il valore della soglia di utilizzo del buffer di uscita della porta, che determina quando il sistema genera il syslog BUFFER_THRESHOLD_EXCEEDED per indicare che l'utilizzo del buffer della slice ASIC ha superato la soglia configurata. A tale scopo, usare il comando di configurazione globale **hardware profile buffer info port-threshold**. Il valore di configurazione predefinito è 90%. Questa configurazione può essere modificata globalmente o per modulo. Di seguito è riportato un esempio di questo comando di configurazione, con il valore 80%.

```
switch# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
switch(config)# hardware profile buffer info port-threshold threshold 80
switch(config)# end
switch# show running-config | include hardware.profile.buffer
hardware profile buffer info port-threshold threshold 80
switch#
```

È possibile modificare l'intervallo minimo tra i syslog BUFFER_THRESHOLD_EXCEEDED generati dallo switch. È inoltre possibile disabilitare completamente il syslog BUFFER_THRESHOLD_EXCEEDED. A tale scopo, usare il comando di configurazione globale **timer syslog-interval info del profilo hardware**. Il valore di configurazione predefinito è 120 secondi. È possibile disattivare completamente il syslog impostando il valore su 0 secondi. Di seguito è riportato un esempio di questo comando di configurazione, con il quale il syslog è interamente disabilitato.

```
switch# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
switch(config)# hardware profile buffer info syslog-interval timer 0
switch(config)# end
switch# show running-config | include hardware.profile.buffer
hardware profile buffer info syslog-interval timer 0
switch#
```

Registri da raccogliere per scenari di congestione della rete

È possibile raccogliere i log mostrati qui da uno switch interessato da uno scenario di congestione della rete per identificare un'interfaccia di uscita congestionata in aggiunta ai comandi elencati in questo documento.

1. L'output del comando **show tech-support details**.
2. L'output del comando **show tech-support usd-all**.
3. L'output del comando **show tech-support ipqos all**.
4. Quando si lavora con uno switch Nexus serie 9500 con schede di linea Cisco Cloud Scale inserite, l'output del comando **show system internal interface counters peak module {x}**, dove {x} è il numero di slot del modulo che ospita l'interfaccia in uscita congestionata.

Monitoraggio dei micro-burst

In caso di congestione o sovrascrittura a intervalli molto brevi (micro-burst), sono necessarie

ulteriori informazioni per ottenere una descrizione accurata degli effetti della sovrascrittura sullo switch.

Gli switch Cisco Nexus serie 9000 dotati di Cisco Cloud Scale ASIC possono monitorare il traffico per rilevare micro-burst che possono causare una congestione temporanea della rete e perdita di traffico nell'ambiente. Per ulteriori informazioni sui micro-burst e su come configurare questa funzione, consultare i documenti mostrati di seguito:

- [Capitolo "Micro-Burst Monitoring" della guida alla configurazione della qualità del servizio di Cisco Nexus serie 9000 NX-OS, versione 10.1\(x\)](#)
- [Capitolo "Micro-Burst Monitoring" della guida alla configurazione della qualità del servizio di Cisco Nexus serie 9000 NX-OS, versione 9.3\(x\)](#)
- [Capitolo "Micro-Burst Monitoring" della guida alla configurazione della qualità del servizio di Cisco Nexus serie 9000 NX-OS, versione 9.2\(x\)](#)
- [Capitolo "Micro-Burst Monitoring" della guida alla configurazione della qualità del servizio di Cisco Nexus serie 9000 NX-OS, versione 7.x](#)

Informazioni su questa traduzione

Cisco ha tradotto questo documento utilizzando una combinazione di tecnologie automatiche e umane per offrire ai nostri utenti in tutto il mondo contenuti di supporto nella propria lingua. Si noti che anche la migliore traduzione automatica non sarà mai accurata come quella fornita da un traduttore professionista. Cisco Systems, Inc. non si assume alcuna responsabilità per l'accuratezza di queste traduzioni e consiglia di consultare sempre il documento originale in inglese (disponibile al link fornito).