Comprendre le stockage syslog et la congestion du tampon TAHUSD du Nexus 9000

Table des matières

Introduction

Conditions préalables

- Exigences
- Composants utilisés

Informations générales

Comprendre l'architecture de mise en mémoire tampon ASIC évolutive du cloud Cisco Nexus 9000

- Calculer l'ID d'instance pour plusieurs ASIC et tranches

Comprendre la sursouscription et les rejets de sortie

- Comprendre le Syslog BUFFER_THRESHOLD_EXCEEDED
- Comprendre le compteur d'interface des rejets en sortie

Exemple de scénario de sursouscription

Étapes suivantes

Additional Information

- Options de configuration Syslog BUFFER_THRESHOLD_EXCEEDED
- Journaux à collecter pour les scénarios d'encombrement réseau
- Surveillance des micro-salves

Informations connexes

Introduction

Ce document décrit la mise en file d'attente et la mise en mémoire tampon sur les commutateurs Cisco Nexus de série 9000 équipés d'un circuit intégré à application spécifique Cisco évolutif qui exécute le logiciel NX-OS.

Conditions préalables

Exigences

Cisco vous recommande de comprendre les bases de la commutation Ethernet sur les réseaux de support partagé et la nécessité de mettre en file d'attente/en mémoire tampon dans ces réseaux. Cisco vous recommande également de comprendre les notions de base de la qualité de service (QoS) et de la mise en mémoire tampon sur les commutateurs Cisco Nexus. Reportez-vous à la

documentation pour plus de détails :

- Guide de configuration de la qualité de service NX-OS de la gamme Cisco Nexus 9000, version 10.1(x)
- Guide de configuration de la qualité de service NX-OS de la gamme Cisco Nexus 9000, version 9.3(x)
- Guide de configuration de la qualité de service NX-OS de la gamme Cisco Nexus 9000, version 9.2(x)
- Guide de configuration de la qualité de service NX-OS de la gamme Cisco Nexus 9000, version 7.x

Composants utilisés

Les informations contenues dans ce document sont basées sur les commutateurs de la gamme Cisco Nexus 9000 avec l'ASIC Cloud Scale exécutant le logiciel NX-OS version 9.3(8).

La procédure décrite dans ce document s'applique uniquement au matériel Nexus 9200/9300 suivant :

- Commutateurs fixes Nexus 9200/9300
  - N9K-C92160YC-X
  - N9K-C92300YC
  - N9K-C92304QC
  - N9K-C92348GC-X
  - N9K-C9236C
  - N9K-C9272Q
  - N9K-C932C
  - N9K-C9364C
  - N9K-C93108TC-EX
  - N9K-C93108TC-EX-24
  - N9K-C93180LC-EX
  - N9K-C93180YC-EX
  - N9K-C93180YC-EX-24
  - N9K-C93108TC-FX
  - N9K-C93108TC-FX-24
  - N9K-C93180YC-FX
  - N9K-C93180YC-FX-24
  - N9K-C9348GC-FXP
  - N9K-C93240YC-FX2
  - N9K-C93216TC-FX2
  - N9K-C9336C-FX2
  - N9K-C9336C-FX2-E
  - N9K-C93360YC-FX2

    - N9K-C93180YC-FX3
    - N9K-C93108TC-FX3P
    - N9K-C93180YC-FX3S
    - N9K-C9316D-GX
    - N9K-C93600CD-GX
    - N9K-C9364C-GX
    - N9K-C9364D-GX2A
    - N9K-C932D-GX2B

- Cartes de ligne de commutation modulaires Nexus 9500
    - N9K-X97160YC-EX
    - N9K-X9732C-EX
    - N9K-X9736C-EX
    - N9K-X97284YC-FX
    - N9K-X9732C-FX
    - N9K-X9788TC-FX
    - N9K-X9716D-GX

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si votre réseau est en ligne, assurez-vous de bien comprendre l’incidence possible des commandes.

Informations générales

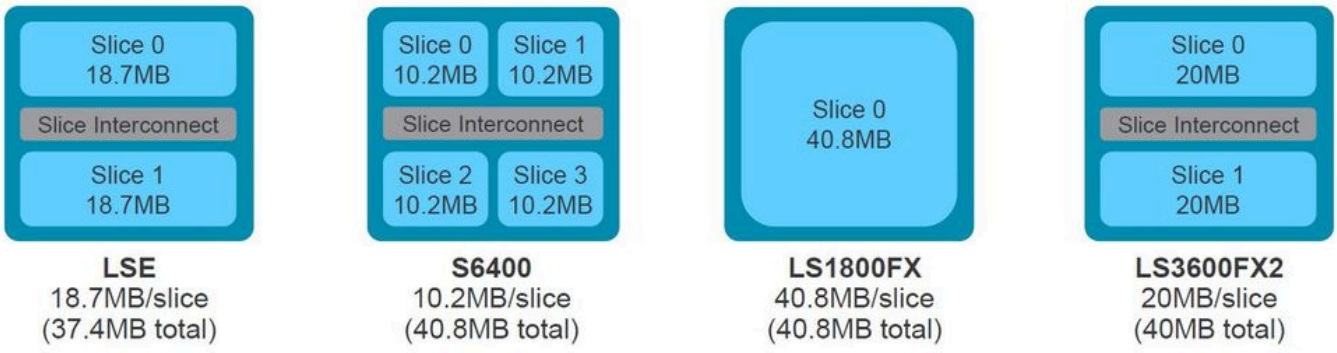
Ce document décrit la mise en file d'attente et la mise en mémoire tampon sur les commutateurs de la gamme Cisco Nexus 9000 équipés de Cisco Cloud Scale ASIC exécutant le logiciel NX-OS.

Comprendre l'architecture de mise en mémoire tampon ASIC évolutive du cloud Cisco Nexus 9000

Les commutateurs Cisco Nexus 9000 avec Cisco Cloud Scale ASIC (Application-Specific Integrated Circuit) implémentent une architecture de tampon de sortie « à mémoire partagée ». Un ASIC est divisé en une ou plusieurs « tranches ». Chaque tranche possède sa propre mémoire tampon et seuls les ports de cette tranche peuvent l’utiliser. Physiquement, chaque tranche est divisée en "cellules", qui représentent des portions de la mémoire tampon. Les tranches sont partitionnées en « pool-groups ». Un certain nombre de cellules sont allouées à chaque pool-group, et elles ne sont pas partagées entre des pools-groups séparés. Chaque groupe de pools possède un ou plusieurs « pools », qui représentent une classe de service (CoS) pour le trafic de monodiffusion ou de multidiffusion. Cela permet à chaque pool-group de garantir des ressources

de tampon pour les types de trafic que le pool-group dessert.

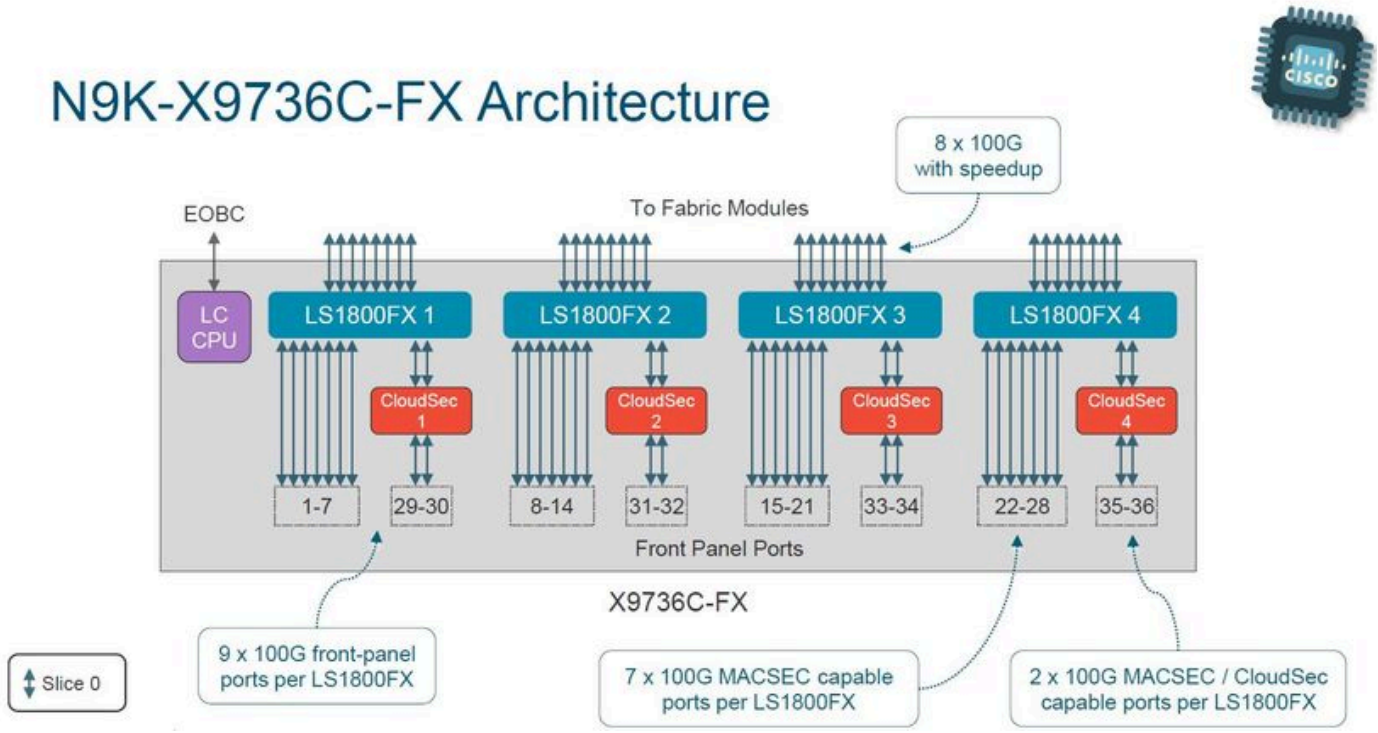
L'image montre comment les différents modèles de Cisco Cloud Scale ASIC sont divisés en tranches. L'image montre également comment chaque tranche alloue une certaine quantité de mémoire tampon à travers les cellules.



Chaque commutateur de la gamme Nexus 9000 et chaque modèle de carte de ligne Nexus 9500 dispose d'un numéro et d'une disposition uniques de circuits ASIC Cisco Cloud Scale qui déterminent les connexions port-ASIC du panneau avant. L'image présente deux exemples qui utilisent la carte de ligne N9K-X9736C-FX et le commutateur N9K-C9336C-FX2.

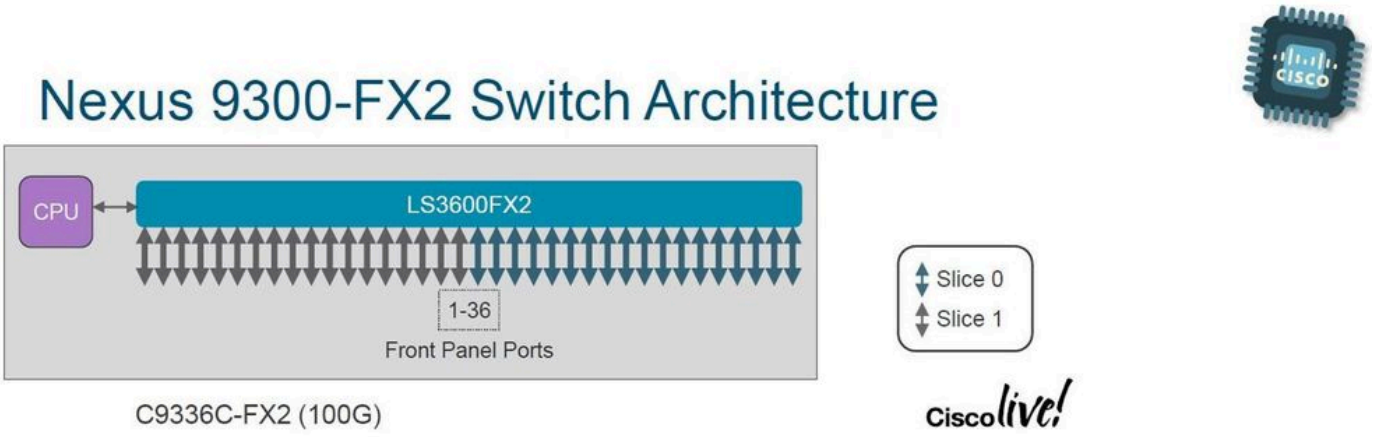
La carte de ligne N9K-C9736C-FX est dotée de 4 circuits ASIC Cisco Cloud Scale LS1800FX avec une tranche par circuit ASIC. En interne, chaque ASIC est appelé une « unité ». Chaque tranche est appelée « instance » et se voit attribuer un entier de base zéro qui identifie de manière unique cette tranche dans le châssis. Il en résulte les permutations :

- L'unité 0, tranche 0, est appelée instance 0
- L'unité 1, tranche 0, est appelée instance 1
- L'unité 2, tranche 0, est appelée instance 2
- L'unité 3, tranche 0, est appelée instance 3



Le commutateur N9K-C9336C-FX2 dispose d'un ASIC Cisco Cloud Scale LS3600FX2 avec deux tranches par ASIC. En interne, chaque ASIC est appelé une « unité ». Chaque tranche est appelée « instance » et se voit attribuer un entier de base zéro qui identifie de manière unique cette tranche dans le châssis. Il en résulte les permutations indiquées ici :

- L'unité 0, tranche 0, est appelée instance 0
- L'unité 0, tranche 1, est appelée instance 1



Chaque carte de ligne et chaque commutateur ont une disposition différente et génèrent différents numéros d'instance. Pour que vous puissiez concevoir votre réseau en fonction de flux de trafic à forte consommation de bande passante, vous devez comprendre la configuration du commutateur ou de la carte de ligne. La commande show interface hardware-mappings peut être utilisée pour

corréler chaque port du panneau avant à une unité (ASIC) et à un numéro de tranche. Dans cet exemple, l'interface Ethernet2/16 d'un commutateur Nexus 9504 avec une carte de ligne N9K-X9736C-FX est insérée dans le logement 2 du châssis et correspond à l'unité 1, tranche 0.

<#root>

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

Calculer l'ID d'instance pour plusieurs ASIC et tranches

Lors de l'interprétation du syslog, l'ID d'instance est calculé en fonction de l'ordre de combinaison d'unités contiguës et de tranches. Par exemple, si un module Nexus 9500 ou un Nexus 9300 TOR (Top-of-Rack) a deux unités (ASIC) et deux tranches par unité, les ID d'instance sont les suivants :

- L'unité 0, tranche 0 est appelée Instance 0
- L'unité 0, tranche 1, est appelée Instance 1
- L'unité 1, tranche 0, est appelée Instance 2
- L'unité 1, tranche 1, est appelée Instance 3

Si un module comporte une unité et quatre tranches, les ID d'instance sont les suivants :

- L'unité 0, tranche 0 est appelée Instance 0
- L'unité 0, tranche 1, est appelée Instance 1
- L'unité 0, tranche 2, est appelée Instance 2
- L'unité 0, tranche 3, est appelée Instance 3

Comprendre la sursouscription et les rejets de sortie

Les interfaces connectées à un réseau Ethernet ne peuvent transmettre qu'un seul paquet à la fois. Lorsque deux paquets doivent sortir d'une interface Ethernet en même temps, l'interface Ethernet transmet un paquet tout en mettant l'autre en mémoire tampon. Une fois le premier paquet transmis, l'interface Ethernet transmet le second paquet à partir du tampon. Lorsque la somme totale du trafic qui doit sortir d'une interface dépasse la bande passante de l'interface, l'interface est considérée comme sursouscrite. Par exemple, si un total de 15 Gbits/s de trafic entre instantanément dans le commutateur et doit sortir d'une interface 10 Gbits/s, l'interface 10 Gbits/s est en sursouscription car elle ne peut pas transmettre 15 Gbits/s de trafic à la fois.

Un commutateur Cisco Nexus 9000 avec un ASIC Cloud Scale gère ce conflit de ressources en mettant en mémoire tampon le trafic dans les mémoires tampon de la tranche ASIC associée à l'interface de sortie. Si la somme totale du trafic qui doit sortir d'une interface dépasse la bande passante de l'interface pendant une période prolongée, les tampons de la tranche ASIC commencent à se remplir de paquets qui doivent sortir de l'interface.

Lorsque les tampons de la tranche ASIC atteignent 90 % d'utilisation, le commutateur génère un syslog semblable à l'exemple suivant :

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

Lorsque les tampons de la tranche ASIC sont complètement pleins, le commutateur abandonne tout trafic supplémentaire qui doit sortir de l'interface jusqu'à ce que l'espace dans les tampons devienne libre. Lorsque le commutateur abandonne ce trafic, il incrémente le compteur Rejets de sortie sur l'interface de sortie.

Le compteur Syslog généré et le compteur de rejets en sortie non nuls sont les deux symptômes d'une interface en sursouscription. Chaque symptôme est étudié plus en détail dans les sous-sections .

Comprendre le Syslog BUFFER_THRESHOLD_EXCEEDED

Exemple du syslog BUFFER_THRESHOLD_EXCEEDED :

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

Ce syslog contient trois éléments d'information clés :

1. Module X : logement de la carte de ligne qui contient l'interface en sursouscription.
2. Instance Y : numéro d'instance affecté à l'ASIC et au tuple de tranches contenant l'interface en sursouscription.
3. Pool-group buffer Z : Seuil de tampon du groupe de pools affecté avant la génération du syslog. Il s'agit d'un pourcentage dérivé des cellules utilisées divisé par le nombre total de cellules tel qu'observé dans le résultat de show hardware internal buffer info pkt-stats lorsqu'il est connecté au module X.

Comprendre le compteur d'interface des rejets en sortie

Le compteur d'interfaces Output Discards indique le nombre de paquets qui ont été abandonnés et qui doivent avoir quitté l'interface, mais qui n'ont pas pu le faire car la mémoire tampon de tranche ASIC était pleine et n'a pas pu accepter de nouveaux paquets. Le compteur Output Discards est visible dans la sortie de show interface et des erreurs show interface counters.

<#root>

```
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 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
    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 oversubcription

    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

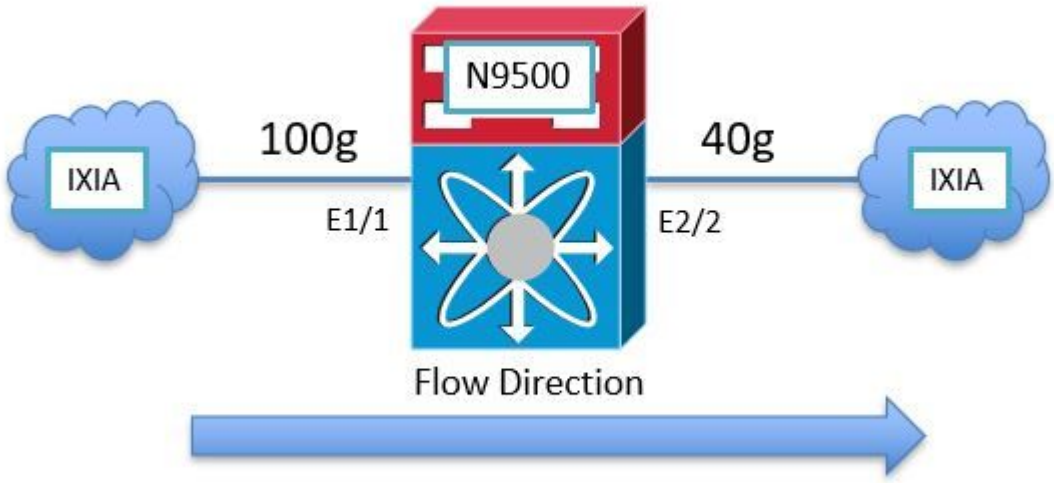

--------------------------------------------------------------------------------
Port         Single-Col  Multi-Col   Late-Col  Exces-Col  Carri-Sen       Runts
--------------------------------------------------------------------------------
Eth1/1                0          0          0          0  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

Exemple de scénario de sursouscription

Considérez ce scénario, où le trafic entre deux générateurs de trafic IXIA traverse un commutateur Nexus 9504 avec deux cartes de ligne N9K-X9736C-FX insérées dans les logements 1 et 2 du châssis. Le trafic de 100 Gbits/s entre dans le commutateur via l'interface Ethernet1/1 de 100 Gbits/s et doit sortir de l'interface Ethernet2/2 de 40 Gbits/s. Par conséquent, Ethernet2/2 est en sursouscription. Exemple de scénario de topologie :



Comme l'ASIC Nexus 9000 Cloud Scale utilise une architecture de mémoire tampon de sortie partagée, vous devez vérifier la mémoire tampon de l'interface de sortie Ethernet2/2 pour voir l'encombrement. Dans cet exemple, la carte de ligne insérée dans le logement 2 est la carte de ligne de sortie. Vous devez donc utiliser la commande attach module 2 avant d'afficher la mémoire tampon matérielle interne avec la commande show hardware internal tah buffer counters. Notez le compteur non nul « Occupancy Drops » pour le groupe de pools Unit 0, Slice 0 et les pools associés, qui indique le nombre de paquets abandonnés parce que la mémoire tampon du groupe de pools est entièrement occupée.

<#root>

```
switch#

attach module 2


module-2#

show hardware internal tah buffer counters



Unit: 0 Slice: 0
```

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

Occupancy drops                 51152554987

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

|-----------------------------------------------------------------------------------------------------
|                                                 Output UC Pool counters
|                             Pool 0     Pool 1     Pool 2     Pool 3     Pool 4     Pool 5     Pool
|-----------------------------------------------------------------------------------------------------
Dynamic Threshold (cells)      93554      93554      93554      93554      93554      93554      9355


Occupancy drops                 51152555398

     0          0          0          0          0          0          0  |
AQM drops                              0          0          0          0          0          0


|-----------------------------------------------------------------------------------------------------
|                                                 Output MC Pool counters
|                             Pool 0     Pool 1     Pool 2     Pool 3     Pool 4     Pool 5     Pool
|-----------------------------------------------------------------------------------------------------
Dynamic Threshold (cells)      93554      93554      93554      93554      93554      93554      9355
Dynamic Threshold (desc)       93554      93554      93554      93554      93554      93554      9355
Dynamic Threshold (inq thr)    64035      64035      64035      64035      64035      64035      6403

Occupancy drops                    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 desciptor drop                  :      0
```

```
Unit: 1 Slice: 0
====================
|-------------------------------------------------------------------------------------------------------
|                                          Output Pool-Group drops
|                              Drop-PG      No-drop      CPU--PG      LCPU-PG      RCPU-PG      SPAN-
|-------------------------------------------------------------------------------------------------------
Occupancy drops                     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
|-------------------------------------------------------------------------------------------------------
Dynamic Threshold (cells)    93554      93554      93554      93554      93554      93554      9355

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

|-------------------------------------------------------------------------------------------------------
|                                            Output MC Pool counters
|                            Pool 0     Pool 1     Pool 2     Pool 3     Pool 4     Pool 5     Pool
|-------------------------------------------------------------------------------------------------------
Dynamic Threshold (cells)    93554      93554      93554      93554      93554      93554      9355
Dynamic Threshold (desc)     93554      93554      93554      93554      93554      93554      9355
Dynamic Threshold (inq thr)  64035      64035      64035      64035      64035      64035      6403

Occupancy drops                 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
          SPAN desciptor drop                   :      0
```

Chaque unité/tuple de tranche ASIC est représenté par une instance unique appelée « instance ». La sortie de la commande show hardware internal buffer info pkt-stats affiche des informations

détaillées sur le groupe de pools encombré (abrégé en "PG") pour chaque instance. La commande affiche également le pic historique/nombre maximal de cellules dans la mémoire tampon qui ont été utilisées.

Enfin, la commande affiche un instantané des identifiants de port ASIC Cloud Scale des ports dont le trafic est mis en mémoire tampon. Un exemple de cette commande :

<#root>

```
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-
|-------------------------------------------------------------------------------------------------

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
Remaining Instant Usage (desc)     93554            0         1500          250         1500         15
Switch Desc Count                  93554            0         1500          250         1500         15

|-------------------------------------------------------------------------------------------------
|                                    Output UC Pool Buffer Utilization (cells/desc)
```

```
|                              Pool 0     Pool 1     Pool 2     Pool 3     Pool 4     Pool 5     Pool
|-------------------------------------------------------------------------------------------------------

Total Instant Usage (cells)      60027

          0          0          0          0          0          0          0  |
Total Instant Usage (desc)           0          0          0          0          0          0

Peak/Max Cells Used              62047

          0          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
|-------------------------------------------------------------------------------------------------------
Total Instant Usage (cells)          0          0          0          0          0          0
Total Instant Usage (desc)           0          0          0          0          0          0
Total Instant Usage (inq cells)      0          0          0          0          0          0
Total Instant Usage (packets)        0          0          0          0          0          0
Peak/Max Cells Used              60399          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]                                                                                          <<< A

          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   |
```

Vous verrez la variation maximale de la commande. Utilisez cette commande pour associer le syslog à un pic potentiel dans un groupe de pools, un pool ou un port particulier.

<#root>

```
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 interface

          count_0       :           0
          count_1       :           0
 [1]
          count_0       :           0
          count_1       :           0
```

La commande show interface hardware-mappings peut être utilisée pour traduire l'identificateur de port ASIC Cloud Scale en port de façade. Dans l'exemple susmentionné, le port ASIC 12 (représenté par la colonne SPort dans la sortie de show interface hardware-mappings) associé à l'unité ASIC 0 sur la tranche/instance 0 a 59 988 cellules occupées de 416 octets chacune. Un exemple de commande show interface hardware-mappings mappe cette interface au port Ethernet2/2 du panneau avant.

<#root>

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

HPortFPort 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
```

En outre, la sursouscription de l'interface Ethernet2/2 avec la mise en file d'attente QoS est abandonnée avec la commande show queuing interface et un exemple :

<#root>

switch#

show queuing interface Ethernet2/2

```
Egress Queuing for Ethernet2/2 [System]
------------------------------------------------------------------------------
QoS-Group# Bandwidth% PrioLevel                Shape                     QLimit
                                     Min          Max        Units
------------------------------------------------------------------------------
      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
```

Vous pouvez vérifier que l'interface de sortie Ethernet2/2 a un compteur d'abandon de sortie non nul avec la commande show interface.

<#root>

```
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 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
    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 oversubcription

    0 Tx pause
```

Étapes suivantes

Si vous observez des rejets de sortie sur un commutateur de la gamme Nexus 9000 avec un ASIC d'évolutivité du cloud, vous pouvez résoudre le problème avec une ou plusieurs des méthodes suivantes :

- Si l'interface dont la sortie est ignorée est une interface unique et n'est pas intégrée dans un port-channel, vous pouvez mettre à niveau la bande passante de l'interface pour réduire la congestion. Par exemple, si une interface de sortie encombrée est une interface 10 Gbits/s et que vous effectuez une mise à niveau vers une interface 25 Gbits/s, 40 Gbits/s ou 100 Gbits/s, elle peut aider à résoudre le problème.
    - En fonction du facteur de forme de l'émetteur-récepteur de l'interface de sortie, vous pouvez mettre à niveau l'émetteur-récepteur (vous pouvez migrer d'un SFP+ 10 Gbit/s inséré dans un CVR-QSFP-SFP10G à l'intérieur d'un port QSFP vers un émetteur-récepteur QSFP 40 Gbit/s natif).
    - Cela peut également être fait si vous migrez la configuration d'interface de sortie encombrée d'un port 10 Gbits/s vers un port 25 Gbits/s, 40 Gbits/s ou 100 Gbits/s.
- Si l'interface dont la sortie est ignorée est une interface unique et n'est pas regroupée dans un port-channel, vous pouvez configurer l'interface encombrée pour qu'elle soit membre d'un port-channel avec une autre interface de même bande passante et réduire l'encombrement.
- Si l'interface dont la sortie est ignorée est une interface port-channel, alors si vous ajoutez des membres au port-channel, cela peut augmenter la bande passante du port-channel global et améliorer le hachage à charge équilibrée pour plusieurs flux de trafic importants.
- Vérifiez si les flux de trafic encombrés entre les hôtes de votre réseau impliquent des interfaces qui diminuent en vitesse (par exemple, le trafic qui entre dans un commutateur via une interface 40 Gbit/s et sort d'un commutateur via une interface 10 Gbit/s). Il peut s'agir d'un goulot d'étranglement qui entraîne une congestion du réseau. Pour éliminer ce goulot d'étranglement, mettez à niveau l'interface à faible débit (par exemple, 10 Gbit/s) vers une interface à débit plus élevé (par exemple, 25 Gbit/s, 40 Gbit/s, etc.), ce qui réduit la congestion du réseau.
- Si vous ne pouvez pas augmenter la bande passante disponible sur l'interface de sortie encombrée, validez la QoS de bout en bout et appliquez les actions de mise en file d'attente appropriées pour votre réseau.
- Si les micro-salves sont une cause potentielle d'encombrement intermittent, consultez la section Surveillance des micro-salves de ce document pour des informations sur la façon de configurer la surveillance des micro-salves.

Additional Information

Cette section du document contient des informations supplémentaires sur les étapes suivantes à suivre lorsque vous rencontrez le syslog BUFFER_THRESHOLD_EXCEEDED, les scénarios d'encombrement/de surabonnement du réseau et les compteurs d'interface d'abandon de sortie incrémentés.

Options de configuration Syslog BUFFER_THRESHOLD_EXCEEDED

Vous pouvez modifier l'intervalle d'interrogation de l'état de la mémoire tampon du système, qui contrôle la fréquence à laquelle le système interroge l'utilisation actuelle des mémoires tampon de tranche ASIC. Pour ce faire, utilisez la commande de configuration globale hardware profile buffer info poll-interval. La valeur de configuration par défaut est de 5 000 millisecondes. Cette configuration peut être modifiée globalement ou sur une base par module. Un exemple de cette commande de configuration est présenté ici, où elle est modifiée à une valeur de 1 000 millisecondes.

<#root>

```
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#
```

Vous pouvez modifier la valeur du seuil d'utilisation du tampon de sortie du port, qui contrôle le moment où le système génère le syslog BUFFER_THRESHOLD_EXCEEDED qui indique que l'utilisation du tampon de tranche ASIC a dépassé le seuil configuré. Pour ce faire, utilisez la commande de configuration globale hardware profile buffer info port-threshold. La valeur de

configuration par défaut est 90 %. Cette configuration peut être modifiée globalement ou sur une base par module. Référez-vous à la commande de configuration où elle est modifiée à la valeur de 80% :

<#root>

```
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#
```

Vous pouvez modifier l'intervalle minimum entre les syslogs BUFFER_THRESHOLD_EXCEEDED générés par le commutateur. Vous pouvez également désactiver le syslog BUFFER_THRESHOLD_EXCEEDED. Pour ce faire, utilisez la commande de configuration globale hardware profile buffer info syslog-interval timer. La valeur de configuration par défaut est de 120 secondes. Le journal système peut être entièrement désactivé en définissant la valeur sur 0 secondes. Référez-vous à la commande de configuration où le syslog est entièrement désactivé :

<#root>

```
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#
```

Journaux à collecter pour les scénarios d'encombrement réseau

Vous pouvez collecter des journaux à partir d'un commutateur affecté par un scénario d'encombrement du réseau pour identifier une interface de sortie encombrée en plus des commandes répertoriées dans ce document.

1. Résultat de la commande show tech-support details.
2. Le résultat de la commande show tech-support usd-all.
3. Résultat de la commande show tech-support ipqos all.
4. Lorsque vous travaillez avec un commutateur de la gamme Nexus 9500 avec des cartes de ligne Cisco Cloud Scale insérées, la sortie de la commande show system internal interface counters peak module {x}, où {x} est le numéro de logement du module hébergeant l'interface de sortie encombrée.

Surveillance des micro-salves

Lorsque la congestion ou la sursouscription se produit à des intervalles très courts (micro-rafale), des informations supplémentaires sont nécessaires pour obtenir une description précise de la manière dont la sursouscription affecte le commutateur.

Les commutateurs de la gamme Cisco Nexus 9000 équipés de l'ASIC Cisco Cloud Scale peuvent surveiller le trafic à la recherche de micro-salves pouvant entraîner un encombrement temporaire du réseau et une perte de trafic dans votre environnement. Reportez-vous à la documentation pour obtenir des informations détaillées sur les micro-rafales et la configuration de cette fonction :

- Chapitre « Micro-Burst Monitoring » du guide de configuration de la qualité de service NX-OS de la gamme Cisco Nexus 9000, version 10.1(x)
- Chapitre « Micro-Burst Monitoring » du guide de configuration de la qualité de service NX-OS de la gamme Cisco Nexus 9000, version 9.3(x)
- Chapitre « Micro-Burst Monitoring » du guide de configuration de la qualité de service NX-OS de la gamme Cisco Nexus 9000, version 9.2(x)
- Chapitre « Micro-Burst Monitoring » du guide de configuration de la qualité de service NX-

OS de la gamme Cisco Nexus 9000, version 7.x

Informations connexes

- Assistance technique de Cisco et téléchargements

À propos de cette traduction
Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.