

Blocages de sortie et QoS de tampon du Nexus 3500

Contenu

[Introduction](#)

[Méthodologie](#)

[Rechercher les pertes de sortie](#)

[Déterminez si les pertes sont de monodiffusion ou de multidiffusion](#)

[Déterminer la mémoire tampon de sortie utilisée](#)

[Vérifier la surveillance de la mémoire tampon active](#)

[Incrémentation active des compteurs](#)

[Bref résultat](#)

[Sortie détaillée](#)

[Générer un journal lorsqu'un seuil est franchi](#)

[ID de bogue Cisco notables](#)

[Forum aux questions](#)

[Annexe - Informations sur les fonctionnalités](#)

[Gestion des tampons](#)

[Planification](#)

[Récepteur lent multidiffusion](#)

[Surveillance de tampon active](#)

[Implémentation matérielle](#)

[Implémentation logicielle](#)

Introduction

Ce document décrit les commandes utilisées afin de dépanner le type de trafic abandonné sur la plate-forme Nexus 3500 et le tampon de sortie (OB) dans lequel ce trafic est abandonné.

Méthodologie

1. [Rechercher les pertes de sortie](#)
2. [Déterminez si les pertes sont de monodiffusion ou de multidiffusion](#)
3. [Déterminer la mémoire tampon de sortie utilisée](#)
4. [Vérifier la surveillance de la mémoire tampon active](#)

Rechercher les pertes de sortie

Vérifiez les statistiques d'interface physique afin de déterminer si le trafic est abandonné dans la direction de sortie. Déterminez si le compteur « rejet de sortie » dans la direction TX augmente et/ou est différent de zéro.

```

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

```

Déterminez si les pertes sont de monodiffusion ou de multidiffusion

Une fois qu'il est déterminé que l'interface abandonne le trafic, entrez la commande **show queuing interface <x/y>** afin de déterminer si le trafic abandonné est de multidiffusion ou de monodiffusion. Dans les versions antérieures à la version 6.0(2)A3(1), la sortie ressemble à :

```

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

```

Dans la version 6.0(2)A3(1) et les versions ultérieures, la sortie ressemble à :

```
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
```

Note: Si le récepteur lent de multidiffusion est configuré pour le port, reportez-vous à la section pour obtenir des informations sur la fonctionnalité, les pertes ne sont pas suivies avec la commande `show queuing interface Eth<x/y>` en raison d'une limitation matérielle. Voir l'ID de bogue Cisco [CSCuj21006](#).

Déterminer la mémoire tampon de sortie utilisée Dans le Nexus 3500, trois pools de tampons sont utilisés dans la direction de sortie. La sortie de la commande `show hardware internal mtc-usd info port-mapping` fournit les informations de mappage.

```
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 première partie des résultats indique que le pool OB 0 est utilisé par les ports avant tels que 45, 46, 47, 48, etc. et OB1 est utilisé par les ports avant 17, 18, etc. La deuxième partie des résultats indique que Eth1/1 est mappé au port OB2 12, Eth1/2 au port OB2 13, etc. Le port en discussion, Eth1/7, est mappé à OB1. Voir la section [Gestion des tampons](#) de ce document pour plus d'informations. **Vérifier la surveillance de la mémoire tampon active** Reportez-vous au livre blanc [Cisco Nexus 3548 Active Buffer Monitoring](#) et à la section de ce document pour plus d'informations sur cette fonctionnalité. **Incrémentation active des compteurs** Si le résultat est rejeté par incrémentation active, activez la surveillance active de tampon (ABM) à l'aide de cette commande. Notez que la commande vous permet de surveiller la monodiffusion ou la multidiffusion, mais pas les deux. Il vous permet également de configurer l'intervalle d'échantillonnage et les valeurs de seuil.

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

Bref résultat Une fois le module ABM activé, vous pouvez afficher les résultats à l'aide de cette commande.

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

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

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

Ces résultats indiquent que 5,376 Mo sur 6 Mo du tampon OB1 ont été utilisés par le trafic de monodiffusion qui a laissé Eth1/7 pendant les 60 dernières secondes. Sortie détaillée

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

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

Legend -

384KB - between 1 and 384KB of shared buffer consumed by port

768KB - between 385 and 768KB of shared buffer consumed by port

307us - estimated max time to drain the buffer at 10Gbps

Active Buffer Monitoring for Ethernet1/7 is: Active

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

Time	384	768	1152	1536	1920	2304	2688	3072	3456	3840	4224	4608	4992	5376	5760
09/30/2013 19:47:01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	250
0 0															
09/30/2013 19:47:00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	252
0 0															
09/30/2013 19:46:59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	253
0 0															
09/30/2013 19:46:58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	250
0 0															
09/30/2013 19:46:57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	250
0 0															
09/30/2013 19:46:56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	250
0 0															
09/30/2013 19:46:55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	251
0 0															
09/30/2013 19:46:54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	251
0 0															
09/30/2013 19:46:53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	250
0 0															
09/30/2013 19:46:52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	253
0 0															
09/30/2013 19:46:51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	249
0 0															
...															

Les informations de chaque ligne sont enregistrées à un deuxième intervalle. Chaque colonne représente l'utilisation de la mémoire tampon. Comme indiqué dans les résultats de la commande, s'il y a une valeur non nulle signalée pour la colonne « 384 », cela signifie que l'utilisation de la mémoire tampon était comprise entre 0 et 384 Ko lorsque l'ABM a interrogé l'utilisation de l'OB. Le nombre non nul est le nombre de fois où l'utilisation a été signalée. Ces résultats indiquent que OB1 a utilisé en moyenne 5,376 Mo entre 249 et 253 fois par seconde au cours des 10 dernières secondes pour Eth1/7. Il faut 4 298 microsecondes (us) pour effacer la mémoire tampon de ce trafic. Générer un journal lorsqu'un seuil est franchi Si le compteur de dépôt et l'utilisation de la mémoire tampon s'incrémentent périodiquement, il est alors possible de définir un seuil et de générer un message de journal lorsque le seuil est franchi.

logging level mtc-usd 5

hardware profile buffer monitor unicast sampling 10 threshold 4608

La commande est configurée pour surveiller le trafic de monodiffusion à un intervalle de 10 nanosecondes et lorsqu'il dépasse 75 % de la mémoire tampon, il génère un journal. Vous pouvez également créer un planificateur afin de collecter des statistiques ABM et la sortie du compteur d'interface toutes les heures et l'ajouter aux fichiers bootflash. Cet exemple concerne le trafic de multidiffusion :

```
hardware profile buffer monitor multicast
```

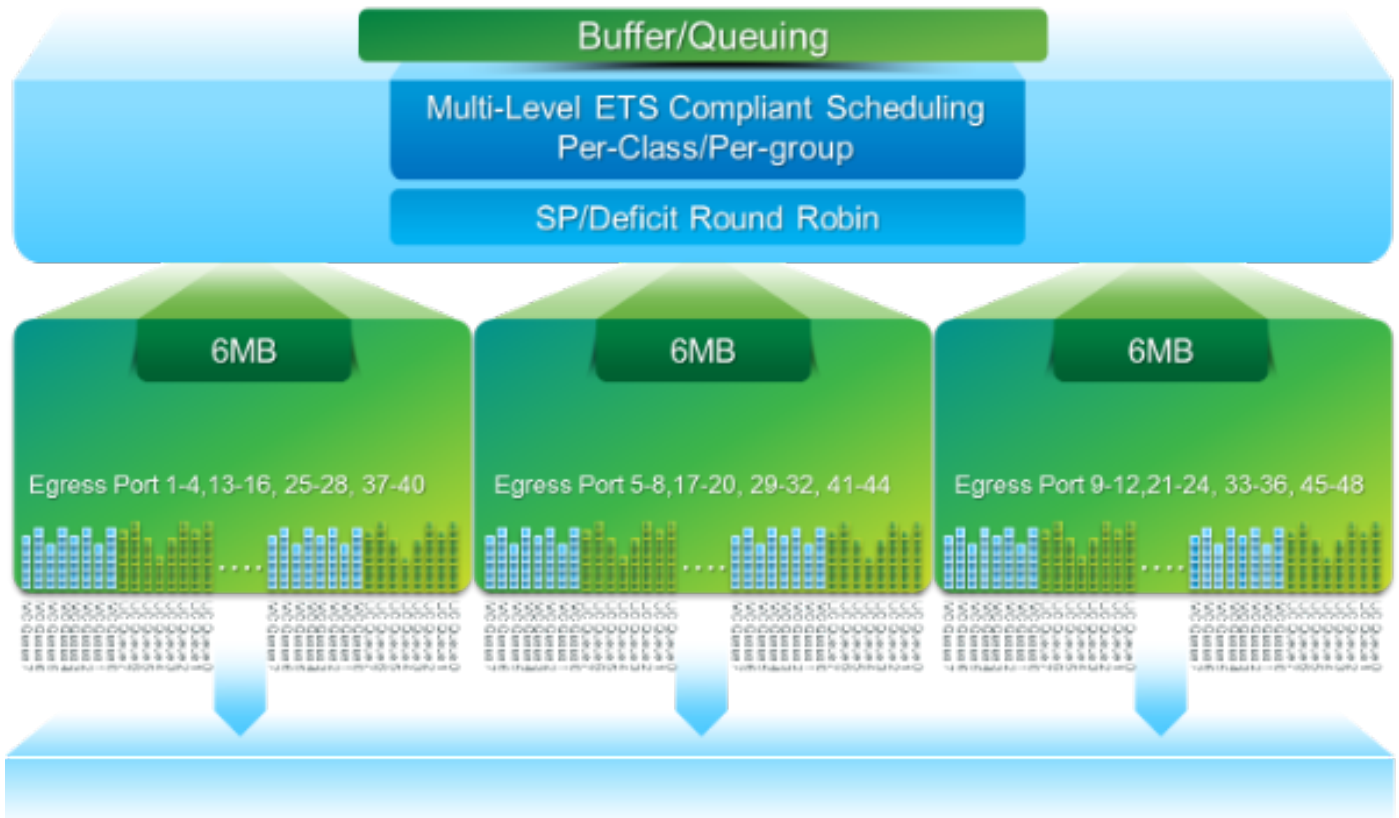
```
feature scheduler
scheduler job name ABM
show hardware profile buffer monitor detail >> ABMDetail.txt
show clock >> ABMBrief.txt
show hardware profile buffer monitor brief >> ABMBrief.txt
show clock >> InterfaceCounters.txt
show interface counters errors >> InterfaceCounters.txt
scheduler schedule name ABM
time start now repeat 1:0
job name ABM
```

ID de bogue Cisco notables

- ID de bogue Cisco [CSCum21350](#) : Les battements de ports rapides font que tous les ports de la même mémoire tampon QoS abandonnent tout le trafic de multidiffusion/diffusion TX. Ceci est corrigé dans la version 6.0(2)A1(1d) et ultérieure.
- ID de bogue Cisco [CSCuq96923](#) : Le bloc de tampon de multidiffusion est bloqué, ce qui entraîne des pertes de multidiffusion/diffusion en sortie. Cette question fait encore l'objet d'une enquête.
- ID de bogue Cisco [CSCva20344](#) : Bloc/verrouillage de mémoire tampon du Nexus 3500 - pas de multidiffusion ou de diffusion TX. Problème non reproductible, potentiellement corrigé dans les versions 6.0(2)U6(7), 6.0(2)A6(8) et 6.0(2)A8(3).
- ID de bogue Cisco [CSCvi93997](#) : bloc de tampon de sortie des commutateurs Cisco Nexus 3500 bloqué. Ceci est corrigé dans les versions 7.0(3)I7(8) et 9.3(3).

Forum aux questions La gestion ABM a-t-elle un impact sur les performances ou la latence ? Non, cette fonctionnalité n'affecte pas la latence ou les performances du périphérique. Quel est l'impact de l'intervalle d'interrogation matériel ABM inférieur ? Par défaut, l'intervalle d'interrogation matériel est de 4 millisecondes. Vous pouvez configurer cette valeur aussi basse que 10 nanosecondes. Il n'y a aucun impact sur les performances ou la latence en raison de l'intervalle d'interrogation matériel plus faible. L'interrogation matérielle par défaut de 4 millisecondes est sélectionnée afin de s'assurer que vous ne débordez pas les compteurs d'histogramme avant que le logiciel n'interroge toutes les secondes. Si vous réduisez l'intervalle d'interrogation du matériel, cela risque de saturer les compteurs matériels à 255 échantillons. Le périphérique ne peut pas gérer une interrogation logicielle inférieure à une seconde, afin de correspondre à l'interrogation matérielle inférieure en raison des restrictions de CPU et de mémoire. Le livre blanc présente l'exemple de l'intervalle d'interrogation matériel le plus faible et

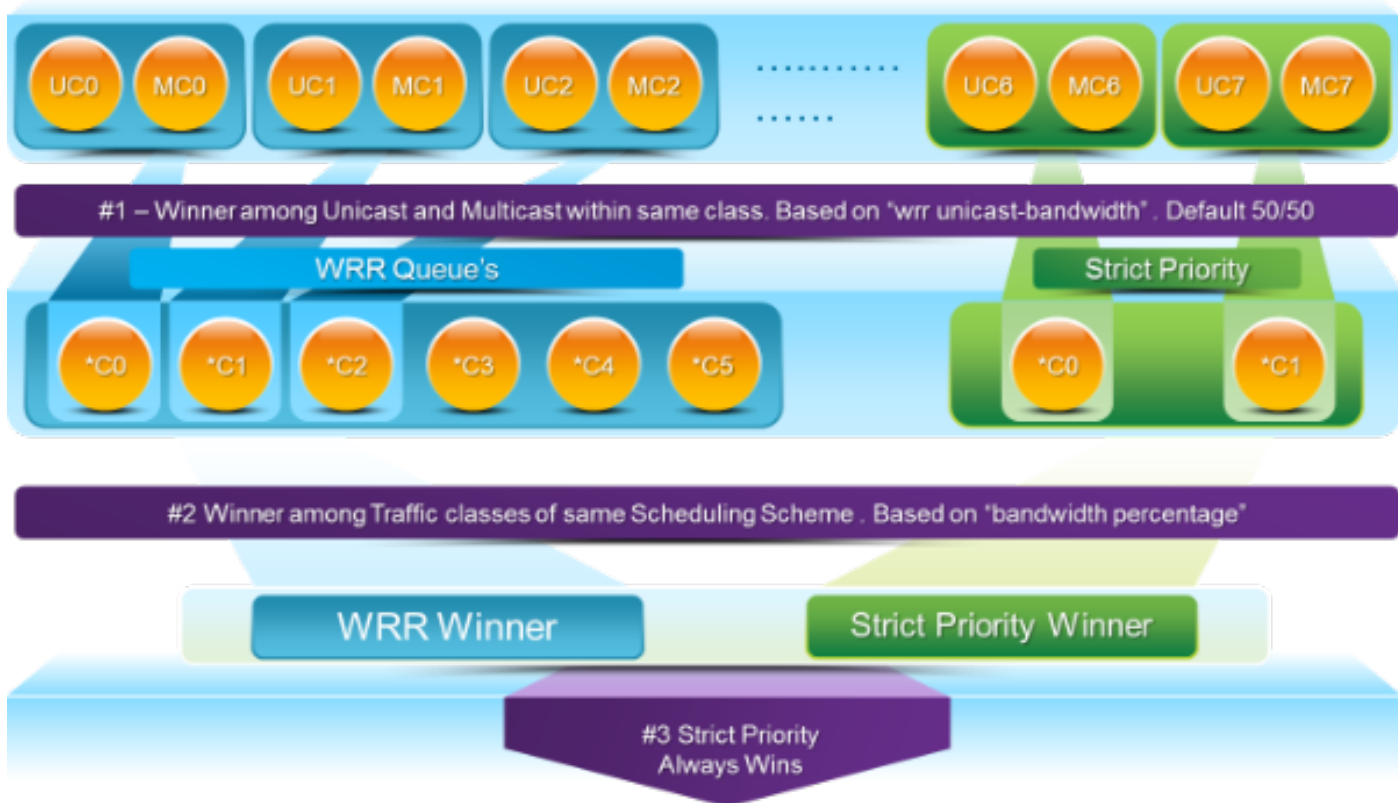
de son cas d'utilisation. **Annexe - Informations sur les fonctionnalités** Gestion des tampons



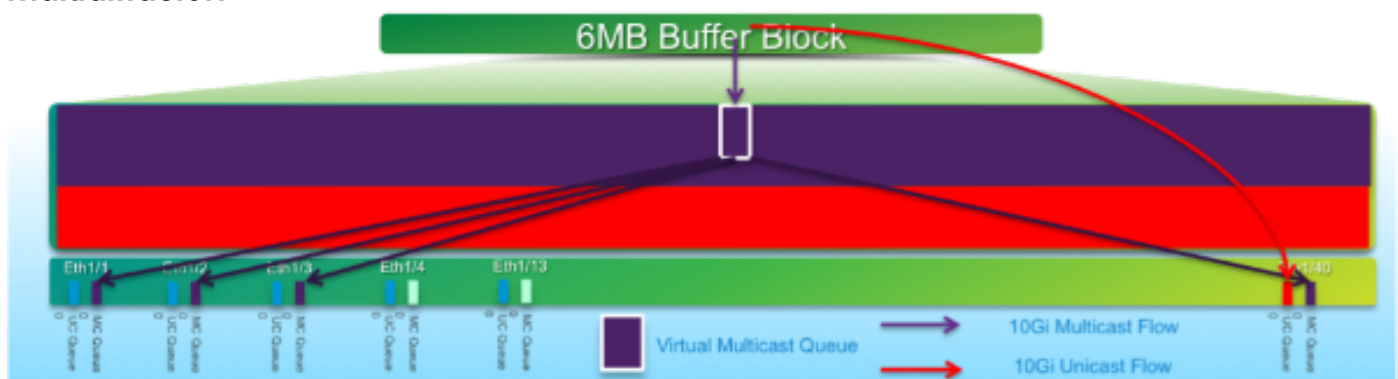
- Mémoire tampon de paquets de 18 Mo partagée par trois blocs OB : ~4 Mo réservés : Taille basée sur l'unité de transmission maximale (MTU) configurée (somme par port de 2 x taille MTU x nombre de groupes QoS activés)~14 Mo partagés : Reste de la mémoire tampon totale~767 Ko d'OB : 0 pour les paquets destinés au processeur
- 6 Mo pour chaque OB sont partagés par un ensemble de 16 ports (commande show hardware internal mtc-usd info port-mapping)

Planification à trois couches :

- Monodiffusion et multidiffusion
- Classes de trafic du même schéma de planification
- Classes de trafic dans le schéma



Récepteur lent multidiffusion



Dans ce schéma :

- Une congestion durable est introduite sur 1 G Eth1/40.
- D'autres récepteurs de multidiffusion (Eth1/1 - 3) sur le bloc de tampon sont affectés en raison du comportement de planification de multidiffusion. *Les récepteurs des autres blocs de mémoire tampon ne sont pas affectés.*
- Le récepteur lent multidiffusion peut être appliqué à e1/40 afin d'éviter la perte de trafic sur les ports non encombrés.
- Le récepteur lent de multidiffusion draine la multidiffusion à un débit de 10 G sur Eth1/40. *Des abandons sont toujours attendus sur le port congestionné.*
- Configuré avec la commande hardware profile multicast slow-récepteur port <x>.

Surveillance de tampon active Reportez-vous au livre blanc [Cisco Nexus 3548 Active Buffer Monitoring](#) pour obtenir une vue d'ensemble de cette fonctionnalité.

- L'ASIC a 18 compartiments et chaque compartiment correspond à une plage d'utilisation de la mémoire tampon (par exemple 0-384 Ko, 385-768 Ko, etc.).
- ASIC interroge l'utilisation de la mémoire tampon pour tous les ports toutes les 4 millisecondes (par défaut). Cet intervalle d'interrogation ASIC est configurable jusqu'à 10 nanosecondes.

- En fonction de l'utilisation de la mémoire tampon pour chaque intervalle d'interrogation matérielle, le compteur de groupement de la plage correspondante est incrémenté. En d'autres termes, si le port 25 consomme 500 Ko de mémoire tampon, le compteur du compartiment n° 2 (385-768 Ko) est incrémenté.
- Ce compteur d'utilisation de la mémoire tampon est maintenu pour chaque interface au format histogramme.
- Chaque compartiment est représenté par 8 bits, de sorte que le compteur dépasse 255 et il est réinitialisé une fois que le logiciel lit les données.

Implémentation logicielle

- Chaque seconde, le logiciel interroge ASIC afin de télécharger et effacer tous les compteurs d'histogramme.
- Ces compteurs d'histogramme sont conservés dans la mémoire pendant 60 minutes avec une granularité d'une seconde.
- Le logiciel s'assure également qu'il copie l'histogramme de tampon dans le bootflash toutes les heures, qui peut être copié dans l'analyseur pour une analyse plus approfondie.
- En effet, cette opération conserve deux heures de données d'histogramme de tampon pour tous les ports, la dernière heure en mémoire et la seconde heure dans le bootflash.