

Dépannage des pannes TX sur Nexus 2232

Contenu

[Introduction](#)

[Conditions préalables](#)

[Conditions requises](#)

[Components Used](#)

[Tampons d'entrée](#)

[Configuration du contrôle de flux](#)

[Causes de la pause TX sur Nexus 2232](#)

[Exemples de tests en laboratoire](#)

[Diagramme du réseau](#)

[Essai 1. Trafic en rafale avec contrôle de flux non activé sur l'hôte](#)

[Essai 2. Trafic en rafale avec contrôle de flux activé sur l'hôte](#)

[Essai 3. Collision de hachage EtherChannel](#)

[Correction](#)

[Conclusions et meilleures pratiques](#)

Introduction

Ce document décrit les informations afin d'aider à dépanner les pauses de transmission (TX) sur les ports HIF (Host Interface) Nexus 2232. Il se concentre sur le trafic dans la direction d'hôte à réseau (H2N) (trafic entrant des serveurs vers le réseau, du sud au nord). Il ne couvre pas les scénarios liés aux flux de trafic N2H (Network to Host).

Ce document est principalement écrit pour Nexus 2232 Fabric Extender (FEX), mais le concept s'applique aux FEX B22 et 2248UPQ.

Conditions préalables

Conditions requises

Cisco vous recommande de prendre connaissance des rubriques suivantes

- Configuration de la gamme Cisco Nexus 2000
- Configuration de la gamme Cisco Nexus 6000

Components Used

Les informations contenues dans ce document sont basées sur les versions de matériel et de logiciel suivantes :

- Cisco Nexus N2K-C2232PP-10GE
- Cisco Nexus 6001

- 7.1(1)N1(1)

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. If your network is live, make sure that you understand the potential impact of any command.

Tampons d'entrée

Le Nexus 2232 dispose de 32 ports HIF (1Host Facing2) 1/10G et de 8 ports NIF (3Network Facing4) 10G.

Avant de vous plonger dans le problème de pause TX, vous devez comprendre les tampons disponibles sur les interfaces FEX. Les tampons affectés à l'interface/qos-group peuvent être vérifiés via cette commande sur le commutateur parent :

```
esc-6001# show queuing interface ethernet 147/1/1
if_slot 79, ifidx 0x1f920000
Ethernet147/1/1 queuing information:
Input buffer allocation:
Qos-group: 0
frh: 8
drop-type: drop
cos: 0 1 2 3 4 5 6
xon      xoff      buffer-size
-----+-----+-----
0        126720    151040
```

<snip>

Comme nous l'avons vu, avec la qualité de service (QoS) par défaut, pour le trafic de classe de rejet (qos-group 0), le HIF FEX a 151040 octets afin de mettre en mémoire tampon le trafic H2N et le seuil XOFF est de 126720 octets.

Configuration du contrôle de flux

Le Nexus 2232 est surabonné à 8:1. Afin d'éviter les pertes de paquets dans la direction H2N en raison de sursouscription et de dépassements de tampon, le Nexus 2232 a le contrôle de flux HIF envoyé par défaut :

```
esc-6001# show run int ethernet 147/1/1 all | inc flow
priority-flow-control mode auto
flowcontrol receive off
flowcontrol send on
esc-6001# show interface ethernet 147/1/1 flowcontrol
```

```
-----
Port          Send FlowControl  Receive FlowControl  RxPause  TxPause
              admin    oper    admin    oper
-----
Eth147/1/1   on      on      off      off      0          0
```

Causes de la pause TX sur Nexus 2232

Lorsque le seuil XOFF de 126720 octets est atteint, le Nexus 2232 envoie une Pause TX vers l'hôte sur le HIF. Les causes les plus courantes sont les suivantes :

1. Le trafic H2N qui entre dans le FEX est très dense, ce qui fait que les tampons d'entrée sont remplis et atteignent le seuil XOFF.

2. La plupart des déploiements FEX utilisent des canaux de port pour agréger plusieurs NIF. La Pause TX est également visible en raison des tampons d'entrée, qui devient saturé en raison d'une collision de hachage EtherChannel sur FEX. Cela se produit lorsque plusieurs ports HIF tentent de sortir d'un seul NIF en raison des résultats etherchannel.

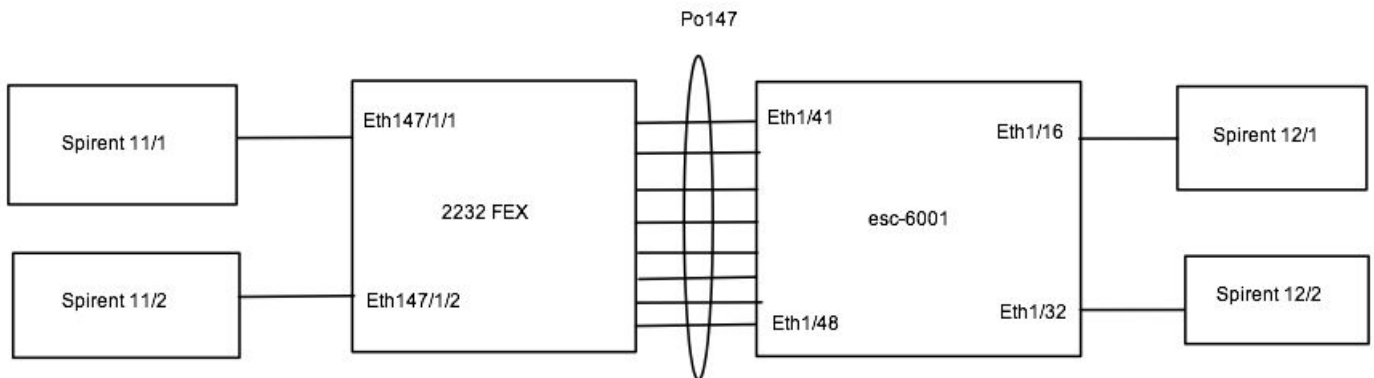
Baisse dans la direction H2N

Afin d'éviter les pertes de paquets, TX Pause est envoyé une fois le seuil XOFF atteint. Cependant, les abandons de trafic H2N peuvent être observés si :

1. les serveurs n'honorent pas la pause ou,
2. ont un délai pour respecter la pause qui provoque l'atteinte du seuil de perte de 151 Ko.

Exemples de tests en laboratoire

Diagramme du réseau



Pour ce test, il existe quatre ports 10G spirent qui agissent comme hôtes, deux sont sur le FEX et deux sont sur le Nexus 6001 parent. Tous les ports se trouvent dans le VLAN 50. Aucun autre port n'est actif sur le FEX ou le parent :

```
esc-6001# show port-channel summary
Flags:  D - Down          P - Up in port-channel (members)
        I - Individual    H - Hot-standby (LACP only)
        s - Suspended     r - Module-removed
        S - Switched      R - Routed
        U - Up (port-channel)
        M - Not in use. Min-links not met
```

```
-----
Group Port-      Type      Protocol  Member Ports
Channel
-----
147   Po147(SU)   Eth       NONE      Eth1/41(P)  Eth1/42(P)  Eth1/43(P)
                                           Eth1/44(P)  Eth1/45(P)  Eth1/46(P)
                                           Eth1/47(P)  Eth1/48(P)
```

```
esc-6001# show fex 147 detail | exc Down
FEX: 147 Description: FEX0147 state: Online
FEX version: 7.1(1)N1(1) [Switch version: 7.1(1)N1(1)]
```

```

FEX Interim version: 7.1(1)N1(1)
Switch Interim version: 7.1(1)N1(1)
Extender Serial: FOT1635R003
Extender Model: N2K-C2232PP-10GE, Part No: 73-12533-05
Card Id: 82, Mac Addr: 20:3a:07:34:5b:02, Num Macs: 64
Module Sw Gen: 21 [Switch Sw Gen: 21]
post level: complete
Pinning-mode: static Max-links: 1
Fabric port for control traffic: Eth1/47
FCoE Admin: false
FCoE Oper: true
FCoE FEX AA Configured: false
Fabric interface state:
  Po147 - Interface Up. State: Active
  Eth1/41 - Interface Up. State: Active
  Eth1/42 - Interface Up. State: Active
  Eth1/43 - Interface Up. State: Active
  Eth1/44 - Interface Up. State: Active
  Eth1/45 - Interface Up. State: Active
  Eth1/46 - Interface Up. State: Active
  Eth1/47 - Interface Up. State: Active
  Eth1/48 - Interface Up. State: Active
Fex Port      State Fabric Port
  Eth147/1/1   Up      Po147
  Eth147/1/9   Up      Po147

```

Logs:

```

04/21/2015 21:58:30.162193: Module register received
04/21/2015 21:58:30.164611: Registration response sent
04/21/2015 21:58:30.196708: create module inserted event.
04/21/2015 21:58:30.197425: Module Online Sequence
04/21/2015 21:58:35.051474: Module Online

```

Essai 1. Trafic en rafale avec contrôle de flux non activé sur l'hôte

Lorsque vous envoyez des rafales de monodiffusion de débit de ligne de 100 000 octets de l'hôte sur Eth147/1/1 (vers Eth1/16) et Eth147/1/9 (do Eth1/32). Chaque flux est un flux unique. Le contrôle de flux est désactivé sur l'hôte (Spirent).

Résultats : Les ports du récepteur ont signalé environ 563 paquets abandonnés pour chaque flux. Comme le contrôle de flux est désactivé sur l'hôte, vous pouvez voir beaucoup plus de Pause TX et aussi une latence élevée (environ 100 microsecondes) :

```

esc-6001# clear counters ; clear qos stat
esc-6001#
esc-6001# show interface ethernet 147/1/1, ethernet 147/1/9 | egrep Ethernet147|pause|unicast
Ethernet147/1/1 is up
  100000 unicast packets  0 multicast packets  0 broadcast packets
  0 Rx pause
  0 unicast packets  578269 multicast packets  0 broadcast packets
  578267 Tx pause
Ethernet147/1/9 is up
  100000 unicast packets  0 multicast packets  0 broadcast packets
  0 Rx pause
  0 unicast packets  578275 multicast packets  0 broadcast packets
  578273 Tx pause
esc-6001# show interface ethernet 147/1/1, eth147/1/9 flowcontrol

```

```

-----
Port          Send FlowControl  Receive FlowControl  RxPause  TxPause
          admin      oper      admin      oper
-----

```

```

Eth147/1/1  on      on      off     off     0      578267
Eth147/1/9  on      on      off     off     0      578273

```

Les pertes signalées par le récepteur sont en fait abandonnées sur le FEX lui-même. Il existe des commandes matérielles internes qui peuvent afficher les pertes, mais cela nécessite que vous compreniez complètement l'architecture interne FEX qui n'est pas comprise dans ce document. Si vous devez vérifier ces compteurs, contactez le TAC pour cet aspect du dépannage.

Essai 2. Trafic en rafale avec contrôle de flux activé sur l'hôte

Lorsque vous envoyez des rafales de monodiffusion à débit unique de 100 000 octets à partir de l'hôte sur Eth147/1/1 (vers Eth1/16) et Eth147/1/9 (do Eth1/32). Chaque flux est un flux unique. Le contrôle de flux est activé sur l'hôte (Spirent).

Résultats :

Les ports du récepteur ne signalent aucune perte. Pause TX minimale et latence moyenne sont d'environ 19 microsecondes :

```

esc-6001# clear counters ; clear qos stat
esc-6001# show interface ethernet 147/1/1, ethernet 147/1/9 | egrep Ethernet147|pause|unicast
Ethernet147/1/1 is up
  100000 unicast packets  0 multicast packets  0 broadcast packets
  0 Rx pause
  0 unicast packets  4743 multicast packets  0 broadcast packets
  4739 Tx pause
Ethernet147/1/9 is up
  100000 unicast packets  0 multicast packets  0 broadcast packets
  0 Rx pause
  0 unicast packets  4703 multicast packets  0 broadcast packets
  4700 Tx pause

```

```

esc-6001# show interface ethernet 147/1/1, eth147/1/9 flowcontrol

```

```

-----
Port          Send FlowControl  Receive FlowControl  RxPause  TxPause
          admin    oper      admin    oper
-----
Eth147/1/1  on      on      off     off     0      4739
Eth147/1/9  on      on      off     off     0      4700

```

Pertes :

Il n'y a pas de perte car l'hôte honore le contrôle de flux envoyé par le FEX.

Essai 3. Collision de hachage EtherChannel

La liaison ascendante entre FEX et parent est un canal de port. Bien que cela dépend du membre choisi dans le port-channel et de l'affluence, la Pause TX peut être vue sur les HIF FEX. Au cours des travaux pratiques, deux ports seulement sont actifs sur le FEX et les 8 liaisons ascendantes sont utilisées dans le port-channel.

Mais pour ce test, avec hachage par défaut, le trafic de l'hôte sur Ethernet 147/1/1 et Ethernet 147/1/9 est haché sur NIF0 qui se connecte à Eth1/41 sur le 6001. Si vous envoyez 98 % de trafic de débit de ligne à partir des hôtes, TX Pause est envoyé sur les deux HIF.

Pour ce test, le contrôle de flux est désactivé sur les hôtes :

esc-6001# show interface ethernet 147/1/1, ethernet 147/1/9 | inc Ethernet14|rate|pause

Ethernet147/1/1 is up

30 seconds input rate 9836009128 bits/sec, 819667 packets/sec

30 seconds output rate 2516922296 bits/sec, 4915863 packets/sec

input rate 9.84 Gbps, 819.67 Kpps; output rate 2.52 Gbps, 4.91 Mpps

0 Rx pause

98376923 Tx pause

Ethernet147/1/9 is up

30 seconds input rate 9836252112 bits/sec, 819687 packets/sec

30 seconds output rate 2516980960 bits/sec, 4915978 packets/sec

input rate 9.84 Gbps, 819.69 Kpps; output rate 2.52 Gbps, 4.91 Mpps

0 Rx pause

98376916 Tx pause

esc-6001# show port-channel traffic interface port-channel 147

ChanId Port Rx-Ucst Tx-Ucst Rx-Mcst Tx-Mcst Rx-Bcst Tx-Bcst

ChanId	Port	Rx-Ucst	Tx-Ucst	Rx-Mcst	Tx-Mcst	Rx-Bcst	Tx-Bcst
147	Eth1/41	99.99%	0.0%	12.50%	12.50%	0.0%	0.0%
147	Eth1/42	0.0%	0.0%	12.50%	12.50%	0.0%	0.0%
147	Eth1/43	0.0%	0.0%	12.50%	12.50%	0.0%	0.0%
147	Eth1/44	0.0%	0.0%	12.50%	12.50%	0.0%	0.0%
147	Eth1/45	0.0%	0.0%	12.50%	12.50%	0.0%	0.0%
147	Eth1/46	0.0%	0.0%	12.50%	12.50%	0.0%	0.0%
147	Eth1/47	0.00%	99.00%	12.50%	12.50%	0.0%	0.0%
147	Eth1/48	0.0%	1.00%	12.50%	12.50%	0.0%	0.0%

esc-6001# attach fex 147

Attaching to FEX 147 ...

To exit type 'exit', to abort type '\$.'

fex-147# dbgexec w

woo> rate

Port	Tx Packets	Tx Rate	Tx Bit	Rx Packets	Rx Rate	Rx Bit	Avg
Pkt Avg Pkt		(pkts/s)	Rate		(pkts/s)	Rate	(Tx)
(Rx) Err							
0-NI8	24	4	11.23Kbps	22	4	16.49Kbps	272
448							
0-NI7	15	3	4.17Kbps	17	3	3.81Kbps	154
120							
0-NI6	1	0	1.76Kbps	1	0	1.08Kbps	1080
656							
0-NI5	1	0	1.76Kbps	1	0	1.08Kbps	1080
656							
0-NI4	1	0	1.76Kbps	1	0	1.08Kbps	1080
656							
0-NI3	1	0	1.76Kbps	1	0	1.08Kbps	1080
656							
0-NI2	1	0	1.76Kbps	1	0	1.08Kbps	1080
656							
0-NI1	1	0	1.76Kbps	1	0	1.08Kbps	1080
656							
0-NI0	4108297	821659	10.05Gbps	1	0	1.08Kbps	1509
656							
0-HI31	1	0	2.28Kbps	1	0	2.28Kbps	1412
1412							
0-HI30	1	0	2.28Kbps	1	0	2.28Kbps	1412
1412							
0-HI29	1	0	2.28Kbps	1	0	2.28Kbps	1412
1412							
0-HI28	1	0	2.28Kbps	1	0	2.28Kbps	1412


```

esc-6001# clear counters ; clear qos stat
esc-6001#
esc-6001# show interface ethernet 147/1/1, ethernet 147/1/9 | inc Ethernet14|rate|pause
Ethernet147/1/1 is up
 30 seconds input rate 4926871976 bits/sec, 410572 packets/sec
 30 seconds output rate 1288637816 bits/sec, 2516870 packets/sec
   input rate 4.93 Gbps, 410.57 Kpps; output rate 1.29 Gbps, 2.52 Mpps
 0 Rx pause
   88129183 Tx pause
Ethernet147/1/9 is up
 30 seconds input rate 4924820632 bits/sec, 410401 packets/sec
 30 seconds output rate 1287225224 bits/sec, 2514111 packets/sec
   input rate 4.92 Gbps, 410.40 Kpps; output rate 1.29 Gbps, 2.51 Mpps
 0 Rx pause
   88069874 Tx pause

```

```

esc-6001# show port-channel traffic interface port-channel 147
ChanId      Port Rx-Ucst Tx-Ucst Rx-Mcst Tx-Mcst Rx-Bcst Tx-Bcst
-----
 147  Eth1/41 99.99%  0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/42 0.0%   0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/43 0.0%   0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/44 0.0%   0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/45 0.0%   0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/46 0.0%   0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/47 0.00% 99.00% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/48 0.0%   1.00% 12.50% 12.50%  0.0%  0.0%

```

```

esc-6001# attach fex 147
Attaching to FEX 147 ...
To exit type 'exit', to abort type '$.'
fex-147# dbgexec w
woo> rate

```

Port	Tx Packets	Tx Rate	Tx Bit	Rx Packets	Rx Rate	Rx Bit	Avg
Pkt Avg Pkt		(pkts/s)	Rate		(pkts/s)	Rate	(Tx)
(Rx) Err							
0-NI8	32	6	19.76Kbps	19	3	16.01Kbps	366
506							
0-NI7	13	2	3.85Kbps	20	4	5.14Kbps	165
140							
0-NI6	1	0	1.76Kbps	2	0	2.16Kbps	1080
656							
0-NI5	1	0	1.76Kbps	2	0	2.16Kbps	1080
656							
0-NI4	1	0	1.76Kbps	2	0	2.16Kbps	1080
656							
0-NI3	1	0	1.76Kbps	2	0	2.16Kbps	1080
656							
0-NI2	1	0	1.76Kbps	2	0	2.16Kbps	1080
656							
0-NI1	1	0	1.76Kbps	2	0	2.16Kbps	1080
656							
0-NI0	4105292	821058	10.04Gbps	2	0	2.16Kbps	1509
656							
0-HI31	1	0	2.28Kbps	1	0	2.28Kbps	1412
1412							
0-HI30	1	0	2.28Kbps	1	0	2.28Kbps	1412
1412							
0-HI29	1	0	2.28Kbps	1	0	2.28Kbps	1412

1412									
0-HI28	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI27	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI26	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI25	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI24	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI23	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI22	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI21	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI20	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI19	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI18	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI17	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI16	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI14	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI13	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI12	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI11	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI10	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI9	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI8	12556848	2511369	1.68Gbps	2049754	409950	4.98Gbps	63		
1500									
0-HI6	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI5	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI4	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI3	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI2	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI1	1	0	2.28Kbps	1	0	2.28Kbps	1412		
1412									
0-HI0	12573036	2514607	1.68Gbps	2051092	410218	4.98Gbps	64		
1499									

Correction

Par défaut, pour le trafic IP, l'équilibrage de charge FEX est basé sur l'adresse MAC/IP source-

destination. Pour les problèmes comme celui-ci, modifiez l'algorithme de hachage afin d'obtenir une meilleure distribution du trafic sur le port-channel de la structure. Utilisez cette méthode si l'équilibrage de charge est inégal. Cette option n'est pas une solution absolue :

```
esc-6001# show port-channel load-balance
```

```
Port Channel Load-Balancing Configuration:
```

```
System: source-dest-ip
```

```
Port Channel Load-Balancing Addresses Used Per-Protocol:
```

```
Non-IP: source-dest-mac
```

```
IP: source-dest-ip source-dest-mac
```

Which hashing algorithm to choose depends on traffic profile. Here are the options available.

```
esc-6001# conf t
```

```
Enter configuration commands, one per line. End with CNTL/Z.
```

```
esc-6001(config)# port-channel load-balance ethernet ?
```

```
destination-ip      Destination IP address
```

```
destination-mac    Destination MAC address
```

```
destination-port   Destination TCP/UDP port
```

```
source-dest-ip     Source & Destination IP address (includes 12)
```

```
source-dest-ip-only Source & Destination IP addresses only
```

```
source-dest-mac   Source & Destination MAC address
```

```
source-dest-port  Source & Destination TCP/UDP port (includes 12 and 13)
```

```
source-dest-port-only Source & Destination TCP/UDP port only
```

```
source-ip         Source IP address
```

```
source-mac       Source MAC address
```

```
source-port      Source TCP/UDP port
```

Conclusions et meilleures pratiques

1. TX Pause est un mécanisme opérationnel normal afin d'éviter les pertes de paquets dans la FEX 2232/2248UPQ/B22.
2. Maximisez le nombre de liaisons ascendantes entre 2232/2248UPQ/B22 FEX et parent. Afin de pouvoir avoir plus de chemins vers le réseau et aussi il aide à avoir des tampons max pour le trafic N2H.
3. Si les liaisons ascendantes entre FEX et parent ne sont pas utilisées uniformément, le changement de hachage de canal de port peut aider.
4. Puisqu'il n'y a pas de commutation locale sur FEX, évitez d'avoir des profils de flux de trafic est-ouest sur les hôtes sur FEX.
5. Évitez les appareils en rafale tels que les périphériques NAS et les châssis lames sur les FEX. Ils doivent être sur le parent.
6. Le nouveau FEX 2348UPQ avec tampon partagé de 32 Mo, dispose d'une mémoire tampon partagée de 1 Mo par HIF pour le trafic H2N afin d'une meilleure absorption des rafales. En outre, avec les liaisons ascendantes NIF 40G, les risques de collisions de hachage/d'encombrement de liaison ascendante sont grandement réduits.