

Limite de multi-alimentation de la distance F1 de FCoE de Nexus 7000

Contenu

[Introduction](#)

[Problème](#)

[Solution](#)

[Matériel/logiciel de mise à jour afin de prendre en charge F2/F2e](#)

Introduction

Ce document décrit quoi faire si vous éprouvez des écarts d'entrée sur la Manche de fibre au-dessus des interfaces de multi-alimentation d'Ethernets (FCoE). Ces problème/document de solution est utile quand des symptômes d'écart sont identifiés sur les interfaces qui interconnectent des centres d'hébergement distants.

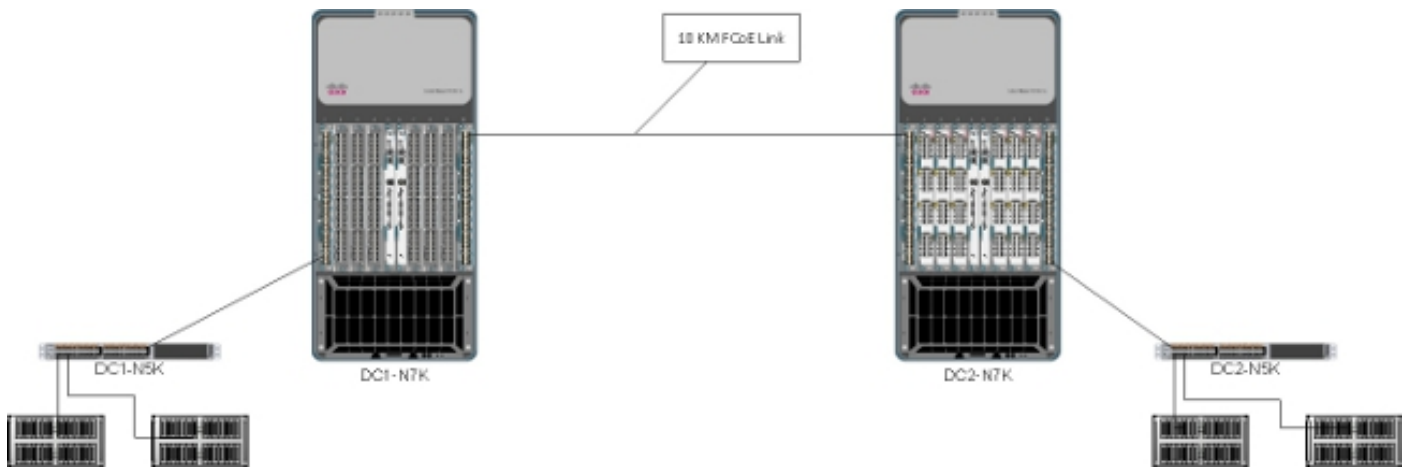
Cet exemple dépeint un vrai scénario de cette question.

La topologie représentée dans l'exemple dépeint deux centres d'hébergement séparés de 10KM. Il y a une extension virtuelle de 10KM FCoE (VE) (connexion multiple entre deux noeuds) l'interface qui connecte DC1 et DC2. Les interfaces de multi-alimentation sont configurées sur les linecards N7K-F132XP-15. Par cette [fiche technique de gamme F1](#), ceci devrait avoir été dans la marge prise en charge.

Au commencement, la fiche technique a indiqué ces IEEE Data Center jetant un pont sur les caractéristiques (le bloc de contrôle de données) :

- contrôle de flux basé sur priorité (PFC) : IEEE P802.1Qbb
- Sélection améliorée de transmission (ETS) : IEEE P802.1Qaz
- Data Center jetant un pont sur l'échange (DCBX)
- Distance sans perte maximum de lien : 20 kilomètres

L'ID de bogue Cisco [CSCts72420](#) a été modifié afin d'adresser la documentation. La ligne en vue de la distance sans perte de lien de 20KM a été retirée.



Problème

Les périphériques EMC VPLEX prennent en charge une caractéristique de réplication de mémoire. Cette réplication synchrone utilisée par scénario. Quand les périphériques EMC VPLEX ont été mis à jour ils sont devenus « hors du sync ». La mise à jour du courrier VPLEX, les périphériques a commencé à répliquer des montants élevés de données au-dessus du lien de multi-alimentation de 10KM FCoE.

Quand la réplication des données a augmenté, ces événements ont transpiré :

1. Le Nexus 5000-DC2 a commencé à envoyer la pause constante au Nexus 7000-DC2.
2. Le Nexus 7000-DC2 a commencé à envoyer la pause au-dessus du lien de 10KM au Nexus 7000-DC1.
3. Le Nexus 7000-DC1 envoyé pause au Nexus 5000-DC1, et ainsi de suite.

Ces événements sont une vue générale du comportement de contrôle de flux prévu de FCoE. Les trames de pause reçues du Nexus 5000-DC2 indiquent des encombrements sur un fin-périphérique. Comme les mémoires tampons d'entrée commencent à remplir, faire une pause des trames écoutez-vous goutte à goutte de nouveau dans la matrice.

La question dans ce scénario est que le Nexus 7000-DC2 a constamment jeté des paquets sur le d'entrée au-dessus du lien de mulithop de 10KM.

```

Ethernet4/1 is up
Dedicated Interface
Hardware: 1000/10000 Ethernet, address: XXXX.XXXX.XXXX (bia XXXX.XXXX.XXXX)
MTU bytes (CoS values): 9216(0-2,4-7) 2112(3)
BW 10000000 Kbit, DLY 10 usec, reliability 255/255, txload 1/255, rxload 1/255
Encapsulation ARPA
Port mode is trunk
full-duplex, 10 Gb/s, media type is 10G
Beacon is turned off
Auto-Negotiation is turned on
Input flow-control is off, output flow-control is off
Rate mode is dedicated
Switchport monitor is off
EtherType is 0x8100
Last link flapped 25week(s) 0day(s)
Last clearing of "show interface" counters 79w2d
30 seconds input rate 296186536 bits/sec, 27891 packets/sec
30 seconds output rate 151677360 bits/sec, 19294 packets/sec
Load-Interval #2: 5 minute (300 seconds)

```

```
input rate 289.58 Mbps, 27.61 Kpps; output rate 165.20 Mbps, 20.05 Kpps
RX
566235497816 unicast packets 2504479 multicast packets 0 broadcast packets
566239834433 input packets 502487779153524 bytes
219280594774 jumbo packets 0 storm suppression packets
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 19312516 input discard
1832141 Rx pause
TX
681040135255 unicast packets 2504251 multicast packets 0 broadcast packets
681046392756 output packets 744942450903588 bytes
333793360248 jumbo packets
0 output error 0 collision 0 deferred 0 late collision
0 lost carrier 0 no carrier 0 babble 0 output discard
3753250 Tx pause
5 interface resets
```

Ceci ne devrait pas se produire pendant que l'interface ci-dessus porte seulement FCoE (le trafic de cos 3). Les écarts d'entrée violent la stratégie QoS de « NO--baisse » pour FCoE. En outre, les écarts dans un environnement de FCoE ont pu mener aux arrêts SCSI, des erreurs, et ainsi de suite.

Solution

Quand un périphérique envoie la pause, l'interface qui génère la trame de pause devrait avoir une file d'attente d'entrée avec un espace de mémoire tampon assez grand pour mettre en mémoire tampon deux fois la distance de lien. C'est parce qu'alors ce que la pause est générée le fil pourrait être plein. Par temps où le périphérique contigu reçoit/traité la trame générée de pause, le fil pourrait être plein de nouveau. Ainsi le périphérique qui génère la pause devrait avoir la capacité de mettre en mémoire tampon deux fois la distance de lien.

Sur le calcul, il pourrait y avoir eu les paquets 100+ en vol au-dessus du lien de 10KM. En raison d'une limite ASIC, le linecard de gamme F1 ne peut pas prendre en charge FCoE sans perte sur un lien de 10KM ou plus grand.

Remarque: La mémoire tampon d'entrée (IB) est utilisée pour aligner des paquets en vol avant la pause. Après que la pause soit envoyée, IB n'est plus utilisé. La mémoire tampon de latence (livre) est utilisée pour aligner des paquets en vol après que la pause est envoyée.

Matériel/logiciel de mise à jour afin de prendre en charge F2/F2e

L'ID de bogue Cisco [CSCua10484](#) a adressé le support sans perte long-courrier de la distance F2. Dans la version 6.1(2) et ultérieures NX-OS, on permet ces modifications de configuration.

Remarque: Dans les versions antérieures du code, on permet la modification mais elle ne la prend pas effet.

L'espace laissé dans l'IB pour attraper des paquets peut être calculé en tant que : PL_STOP - PL_PAUSE. Par défaut les valeurs PL_STOP et HWM (PL_PAUSE) sont identiques.


```
Switch(config)# int e4/1
Switch(config-if)# service-policy type queuing input 7I_4q-7e-in
Switch(config-if)# show run int e4/1
```

```
!Command: show running-config interface Ethernet4/1
!Time: Sun Mar 2 21:03:07 2014
```

```
version 6.1(4)
```

```
interface Ethernet4/1
switchport
switchport mode trunk
switchport trunk allowed vlan 1,2990
load-interval counter 2 30
  service-policy type queuing input 7I_4q-7e-in
no shutdown
```

Maintenant, notez que la valeur PL_STOP est plus grande que le seuil supérieur (HWM). Ainsi, on permet une plus grande capacité de mise en mémoire tampon pour IB.

```
module-4# show hardware internal mac port 1 qos configuration | begin IB | end EB
IB
Port page limit : 3584 (1376256 Bytes)
VL# HWM pages(bytes) LWM pages(bytes) Used PL_STOP(HWM & LWM)
pages THR
0 15 ( 5760) 9 ( 3456) 0 15 9
1 2 ( 768) 1 ( 384) 0 2 1
2 15 ( 5760) 9 ( 3456) 0 15 9
3 1161 ( 445824) 1137 ( 436608) 0 3521 1137
4 2 ( 768) 1 ( 384) 0 2 1
5 3 ( 1152) 0 ( 0) 0 3 0
6 2 ( 768) 1 ( 384) 0 2 1
7 2 ( 768) 1 ( 384) 0 2 1
Credited DWRR WT: 216 (0xd8) Uncredited DWRR WT: 144 (0x90)
DWRR honor UC = FALSE
Leak Lo weight = 0xd8, enabled = FALSE
EB
```

Dans l'exemple, l'espace laissé dans IB = 3521 pages - 1161 pages de = => 2360 pages 906,240 octets.

Remarque: Ce contournement est acceptable quand SEULEMENT le trafic de FCoE est passé au-dessus de l'interface de connexion multiple entre deux noeuds de FCoE. Si le trafic de données est aussi bien passé, entrez en contact avec le centre d'assistance technique Cisco (TAC) pour l'assistance.

Ou

Si disponible, utilisez la Manche indigène de fibre (FC) entre les sites. Cette solution exige le multiplexeur brut de Division de longueur d'onde/intervention dense du multiplexeur de Division de longueur d'onde (CWDM/DWDM) ou la fibre foncée, dépendants sur la distance exigée.