

Limitación de distancia Multihop F1 de FCoE del nexo 7000

Contenido

[Introducción](#)

[Problema](#)

[Solución](#)

[Equipo y programas de computación de la actualización para soportar F2/F2e](#)

Introducción

Este documento describe qué hacer si usted experimenta los descartes de la entrada en el Fibre Channel sobre las interfaces Multihop de los Ethernetes (FCoE). Este problema/documento de soluciones es útil cuando los síntomas del descarte se identifican en las interfaces que interconectan los datacenters remotos.

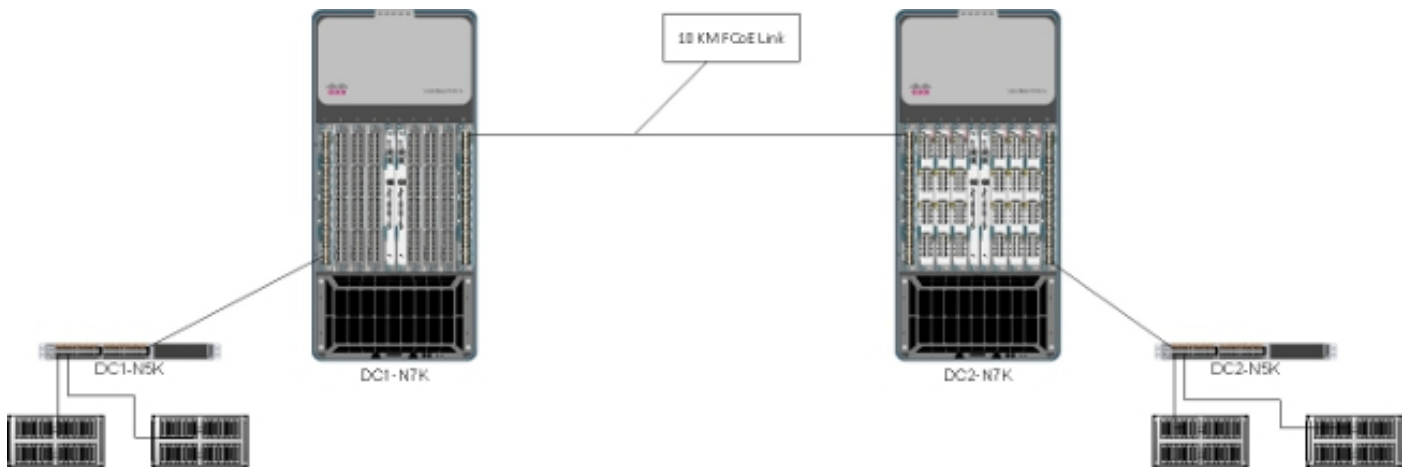
Este ejemplo representa un escenario en la realidad de este problema.

La topología mostrada en el ejemplo representa dos datacenters separados por el 10KM. Hay una extensión virtual del 10KM FCoE (VE) la interfaz (del multihop) que conecta el DC1 y DC2. Las interfaces multihop se configuran en el linecards N7K-F132XP-15. Por esta [ficha técnica de las F1 Series](#), esto debe haber estado dentro del rango admitido.

Inicialmente, la ficha técnica indicó estas características del bridging del centro de datos de IEEE (bcd):

- Control de flujo basado en prioridades (PFC): IEEE P802.1Qbb
- Selección aumentada de la transmisión (ET): IEEE P802.1Qaz
- Intercambio del bridging del centro de datos (DCBX)
- Distancia del link sin pérdidas máxima: 20 kilómetros

El Id. de bug Cisco [CSCts72420](#) fue modificado para dirigir la documentación. La línea con respecto a la distancia del link sin pérdidas de los 20KM fue quitada.



Problema

Los dispositivos EMC VPLEX soportan una característica de la replicación del almacenamiento. Esta replicación síncrona usada escenario. Cuando los dispositivos EMC VPLEX fueron actualizados se convirtieron “fuera de sincronizan”. La actualización del poste VPLEX, los dispositivos comenzó a replicar las grandes cantidades de datos sobre el link multihop del 10KM FCoE.

Cuando la replicación de los datos aumentó, estos eventos transpiraron:

1. El nexa 5000-DC2 comenzó a enviar la pausa constante al nexa 7000-DC2.
2. El nexa 7000-DC2 comenzó a enviar la pausa sobre el link del 10KM al nexa 7000-DC1.
3. El nexa 7000-DC1 enviado pausa al nexa 5000-DC1, y así sucesivamente.

Estos eventos son una vista de alto nivel del comportamiento de control de flujo previsto de FCoE. Las tramas de pausa recibidas del nexa 5000-DC2 indican las congestiones en un dispositivo final. Mientras que memorias intermedias de ingreso comienzan a llenar, las tramas de pausa gotean nuevamente dentro de la tela.

El problema en este escenario es que el nexa 7000-DC2 desechó constantemente los paquetes en el ingreso sobre el link del multihop del 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
```

Esto no debe suceder mientras que la interfaz antedicha lleva solamente FCoE (tráfico de CoS 3). Los descartes de la entrada violan el “ninguno-descenso” política de calidad de servicio (QoS) para FCoE. Además, los descartes en un entorno de FCoE podían llevar a SCSI los abortos, los errores, y así sucesivamente.

Solución

Cuando un dispositivo envía la pausa, la interfaz que genera la trama de pausa debe tener una cola del ingreso con un espacio del búfer bastante grande para mitigar dos veces la distancia del link. Esto es porque en ese entonces ése que la pausa se genera el alambre pudo ser lleno. Por el tiempo que el dispositivo adyacente recibe/que procesa la trama de pausa generada, el alambre pudo ser lleno otra vez. Así el dispositivo que genera la pausa debe tener la capacidad de mitigar dos veces la distancia del link.

Sobre el cálculo, habría podido haber los paquetes 100+ en vuelo sobre el link del 10KM. Debido a una limitación ASIC, el linecard de las F1 Series no puede soportar FCoE sin pérdidas en un link del 10KM o mayor.

Note: Memoria intermedia de ingreso (IB) se utiliza para hacer cola los paquetes en vuelo antes de la pausa. Después de que se envíe la pausa, IB es no se utilizan más. El buffer del tiempo de espera (LB) se utiliza para hacer cola los paquetes en vuelo después de que se envía la pausa.

Equipo y programas de computación de la actualización para soportar F2/F2e

El Id. de bug Cisco [CSCua10484](#) dirigió el soporte sin pérdidas de larga distancia de la distancia F2. En la versión NX-OS 6.1(2) y posterior, se permiten estos cambios de configuración.

Note: En las versiones anteriores del código, se permite el cambio pero no toma el efecto.

El espacio dejado en el IB para coger los paquetes se puede calcular como: PL_STOP - PL_PAUSE. Por abandono los valores PL_STOP y del HWM (PL_PAUSE) son lo mismo.

En el almacenamiento VDC

```
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
```

Ahora, note que el valor PL_STOP es mayor que el punto más alto (HWM). Así, una mayor capacidad que mitiga se permite para 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
```

En el ejemplo, el espacio dejado en IB = 3521 páginas - 1161 páginas de = => 2360 páginas 906,240 bytes.

Note: Esta solución alternativa es aceptable cuando SOLAMENTE el tráfico de FCoE se pasa sobre la interfaz multihop de FCoE. Si el tráfico de datos se pasa también, entre en contacto el Centro de Asistencia Técnica de Cisco (TAC) para la ayuda.

O

Si está disponible, utilice el Fibre Channel nativo (FC) entre los sitios. Esta solución requiere la intervención gruesa o la fibra oscura del multiplexor de la división de longitud de onda/del multiplexor de la división de longitud de onda densa (CWDM/DWDM), dependiente sobre la distancia requerida.