

# Nexus 7000 FCoE Multihop F1 거리 제한

## 목차

[소개](#)

[문제](#)

[솔루션](#)

[F2/F2e 지원을 위해 하드웨어/소프트웨어 업그레이드](#)

## 소개

이 문서에서는 FCoE(Fibre Channel over Ethernet) 다중 hop 인터페이스에서 입력 폐기 현상이 발생하는 경우 수행할 작업에 대해 설명합니다. 이 문제/솔루션 문서는 원격 데이터 센터를 상호 연결하는 인터페이스에서 폐기 증상이 확인될 때 유용합니다.

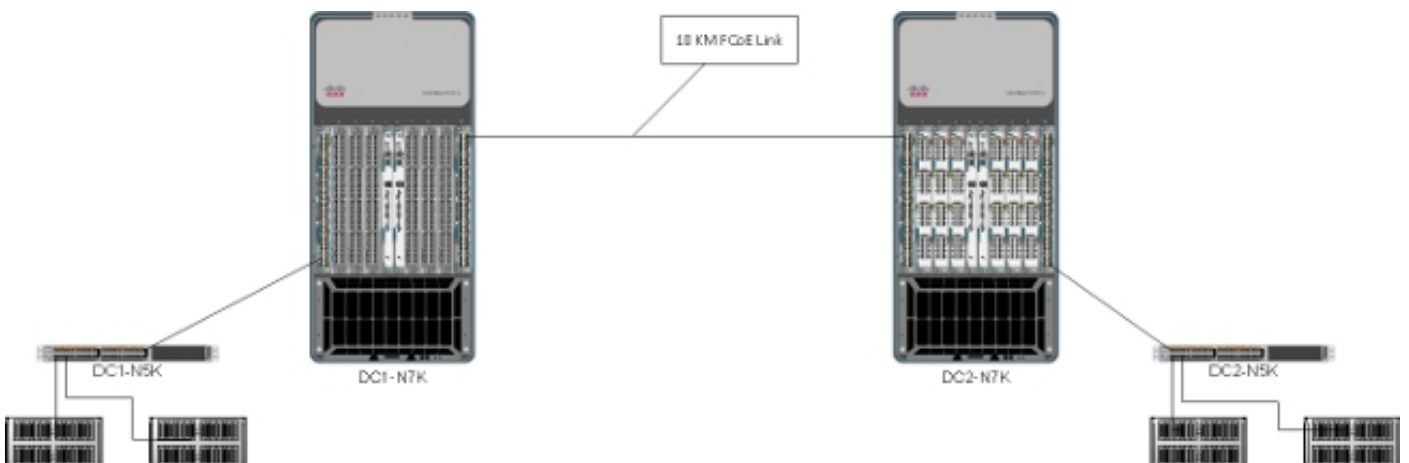
이 예에서는 이 문제의 실제 시나리오를 보여 줍니다.

예시에 표시된 토폴로지는 두 개의 데이터 센터를 10KM로 구분합니다. DC1 및 DC2를 연결하는 10KM FCoE VE(Virtual Expansion)(멀티hop) 인터페이스가 있습니다. 멀티hop 인터페이스는 N7K-F132XP-15 라인 카드에 구성됩니다. 이 [F1 시리즈 데이터시트에](#) 따라 지원되는 범위 내에 있어야 합니다.

처음에 데이터시트는 다음과 같은 IEEE DCB(Data Center Bridging) 기능을 표시했습니다.

- PFC(Priority-based flow control):IEEE P802.1Qbb
- 향상된 전송 선택(ETS):IEEE P802.1Qaz
- DCBX(Data Center Bridging Exchange)
- 최대 무손실 링크 거리:20킬로미터

Cisco 버그 ID [CSCts72420](#)이 문서를 해결하기 위해 수정되었습니다. 20KM의 무손실 링크 거리와 관련된 회선이 제거되었습니다.



# 문제

EMC VPLEX 디바이스는 스토리지 복제 기능을 지원합니다. 이 시나리오에서는 동기식 복제를 사용했습니다. EMC VPLEX 디바이스를 업그레이드하면 '동기화되지 않음'이 됩니다. VPLEX 업그레이드 후 디바이스는 10KM FCoE 멀티홉을 통해 대량의 데이터를 복제하기 시작했습니다.

데이터 복제가 증가하면 다음과 같은 이벤트가 발생합니다.

1. Nexus 5000-DC2는 Nexus 7000-DC2에 지속적인 일시 중지를 보내기 시작했습니다.
2. Nexus 7000-DC2는 10KM 링크를 통해 Nexus 7000-DC1에 일시 중지를 보내기 시작했습니다.
3. Nexus 7000-DC1은 Nexus 5000-DC1에 일시 중지를 전송했습니다.

이러한 이벤트는 예상되는 FCoE 흐름 제어 동작에 대한 대략적인 보기입니다. Nexus 5000-DC2에서 수신한 일시 중지 프레임은 엔드 디바이스의 혼동을 나타냅니다. 인그레스 버퍼가 채워지기 시작하면, 일시 중지 프레임이 패브릭으로 다시 흐릅니다.

이 시나리오의 문제는 Nexus 7000-DC2가 10KM mutisop 링크를 통해 인그레스(ingress)에서 지속적으로 패킷을 폐기했다는 것입니다.

```
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
```

위의 인터페이스에서는 FCoE(CoS 3) 트래픽만 전달하므로 이 오류가 발생하지 않아야 합니다. 입

력 삭제는 FCoE에 대한 'no-drop' QoS 정책을 위반합니다. 또한 FCoE 환경에서 폐기 시 SCSI 중단, 오류 등이 발생할 수 있습니다.

## 솔루션

디바이스가 일시 중지를 전송할 때 일시 중지 프레임을 생성하는 인터페이스에는 링크 거리의 2배에 버퍼링할 수 있을 만큼 큰 버퍼 공간이 있는 인그레스 대기열이 있어야 합니다. 일시정지가 생성되면 와이어가 가득 찰 수 있기 때문입니다. 인접한 디바이스에서 생성된 일시 중지 프레임을 수신/처리할 때까지, 와이어가 다시 가득 찰 수 있습니다. 따라서 일시 중지를 생성하는 디바이스는 링크 거리의 2배를 버퍼링할 수 있어야 합니다.

계산하면 10KM 링크를 통해 100개 이상의 패킷이 비행할 수 있었습니다. ASIC 제한으로 인해 F1 시리즈 라인 카드는 10KM 이상의 링크에서 무손실 FCoE를 지원할 수 없습니다.

**참고:** IB(Ingress Buffer)는 일시 중지 전에 전송 중인 패킷을 대기열에 넣는 데 사용됩니다. 일시 중지를 전송하면 IB는 더 이상 사용되지 않습니다. LB(Latency Buffer)는 일시 중지를 보낸 후 전송 중인 패킷을 큐에 넣는 데 사용됩니다.

## F2/F2e 지원을 위해 하드웨어/소프트웨어 업그레이드

Cisco 버그 ID [CSCua10484](#)는 F2 장거리 무손실 거리 지원을 처리했습니다. NX-OS 릴리스 6.1(2) 이상에서는 이러한 컨피그레이션을 변경할 수 있습니다.

**참고:** 이전 버전의 코드에서는 변경이 허용되지만 적용되지 않습니다.

IB에 남아 있는 패킷 캐치할 공간은 다음과 같이 계산할 수 있습니다. PL\_STOP - PL\_PAUSE. 기본적으로 PL\_STOP 및 HWM(PL\_PAUSE) 값은 동일합니다.

```
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 1107 ( 425088) 1059 ( 406656) 0 1107 1059
1 2 ( 768) 1 ( 384) 0 2 1
2 1107 ( 425088) 1059 ( 406656) 0 1107 1059
3 1053 ( 404352) 1029 ( 395136) 0 1053 1029
4 2 ( 768) 1 ( 384) 0 2 1
5 231 ( 88704) 159 ( 61056) 0 231 159
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
```

더 큰 버퍼를 CoS(no-drop Class of Service)에 할당하여 더 먼 거리를 지원하도록 이러한 값을 수정할 수 있습니다. 이를 완료하려면 'default-4q-7e-in-policy' QoS(Quality of Service) 정책 맵을 복제합니다.



```

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

```

이 예에서는 IB = 3521페이지 - 1161페이지 = 2360페이지 => 906,240바이트에 남아 있는 공간을 보여 줍니다.

**참고:**FCoE 트래픽만 FCoE 다중 hop 인터페이스를 통해 전달되는 경우 이 해결 방법을 사용할 수 있습니다.데이터 트래픽도 통과한 경우 Cisco TAC(Technical Assistance Center)에 문의하십시오.

또는

사용 가능한 경우 사이트 간에 네이티브 FC(Fibre Channel)를 사용합니다.이 솔루션에는 필요한 거리에 따라 CWDM/DWDM(Caular Wavelength Dividation Multiplexer/Dense Wavelength Dividon Multiplexer) 개입 또는 다크 파이버가 필요합니다.