

# Nexus 5000/6000 Series에서 FEX 성능 문제 해결

## 목차

[소개](#)

[배경 정보](#)

[CLI 탐색](#)

[FEX에 연결](#)

[디버그 실행 모드 입력](#)

[디버그 실행 모드 종료](#)

[FEX 종료](#)

[용어](#)

[호스트 인터페이스\(HI\)](#)

[네트워크 인터페이스\(NI\)](#)

[FEX 패브릭 포트](#)

[FEX ASIC 이름](#)

[전면 포트 매핑](#)

[N2K-C2148T-1GE](#)

[N2K-C2224TP-1GE / N2K-C2248TP-1GE](#)

[N2K-C2232PP-10GE / N2K-C2232TM-10GE](#)

[N2K-C2248TP-E-1G](#)

[N2K-C2248PQ-10GE 및 N2K-C2348UPQ-10GE](#)

[SFP 확인](#)

[손실 찾기](#)

[HI 포트 카운터 보기](#)

[NI 포트 카운터 보기](#)

[기록 삭제 보기](#)

[최근 삭제 및 인터럽트 보기](#)

[실시간으로 포트 트래픽 속도 보기](#)

[손실 완화](#)

[서버 위치 변경](#)

[추가 업링크 추가](#)

[HI 버퍼 공유](#)

[Nexus 6000 FEX 로드 밸런스 향상](#)

## 소개

이 문서에서는 Nexus 5000 또는 6000 Series 스위치에 연결할 수 있는 FEX(Fabric Extender)의 성능 문제를 해결하는 방법에 대해 설명합니다.

**참고:** 이 문서에 도입된 명령은 중단이 아닙니다. Nexus 2000 스위치가 5000 또는 6000 시리즈 스위치에 연결되어 있어야 합니다.

# 배경 정보

## CLI 탐색

### FEX에 연결

FEX 명령행에서 show 명령을 실행하려면 FEX에 연결합니다.

```
Nexus# 어태치 fex fex  
fex>
```

### 디버그 실행 모드 입력

고급 명령을 실행하고 FEX 기본 이름을 지정하려면 FEX에서 디버그 모드를 입력합니다.FEX 기본 이름은 표 1.을 참조하십시오.

```
fex# dbgexec [prt/우/red/pri]
```

### 디버그 실행 모드 종료

디버그 실행 모드를 종료하려면 Ctrl+C 키보드 시퀀스를 사용합니다.

```
fex> [Ctrl+C]
```

### FEX 종료

fex를 종료하려면 명령 exit를 사용합니다.

```
fex# 종료
```

## 용어

### 호스트 인터페이스(HI)

FEX의 서버를 접하는 포트입니다.일반적으로 전면 포트라고 합니다.FEX의 모든 전면 포트에는 HI 번호가 있습니다.이 번호는 일반적으로 포트 번호와 달라지지만 포트를 참조하기 위해 명령을 트러블슈팅하는 데 사용됩니다.각 ASIC은 전면 포트를 다르게 표시합니다.

### 네트워크 인터페이스(NI)

NI는 상위 스위치에 다시 연결되는 FEX의 FEX 제어 포트입니다.이러한 링크는 네트워크 업링크라고도 합니다.또한 모델에 따라 고유한 NI 번호가 있습니다.

### FEX 패브릭 포트

이러한 포트는 FEX에 대한 고유한 링크의 상위 스위치입니다. 이러한 포트는 `switchport mode fex-fabric` 및 `fex association` 명령으로 구성됩니다.

## FEX ASIC 이름

각 FEX는 다른 ASIC로 설계되었습니다. ASIC 이름의 약어는 디버깅 모드에서 명령을 실행하는 데 사용됩니다.

FEX의 대부분의 모델에는 ASIC가 1개 있지만 2148에는 6개, 각각 8개의 전면 포트가 있습니다. 이를 트러블슈팅 명령에서 `rmon`이라고 합니다.

ASIC 이름 및 관련 재설정이 참조용으로 나열됩니다.

### 표 1.

FEX 모델	ASIC 이름	취소
N2K-C2148T-1GE	삼나무	rw
N2K-C2224TP-1GE	포올라	포트
N2K-C2248TP-1GE		
N2K-C2232PP-10GE	목판	우
N2K-C2232TM-10GE		
N2K-C2248TP-E-1GE	프린스턴	운영
B22	목판	우
N2K-C2232TM-E-10GE	목판	우
N2K-C2248PQ-10GE	우드사이드/벨몬트	우
N2K-C2348UPQ-10GE	티부론	티비

## 전면 포트 매핑

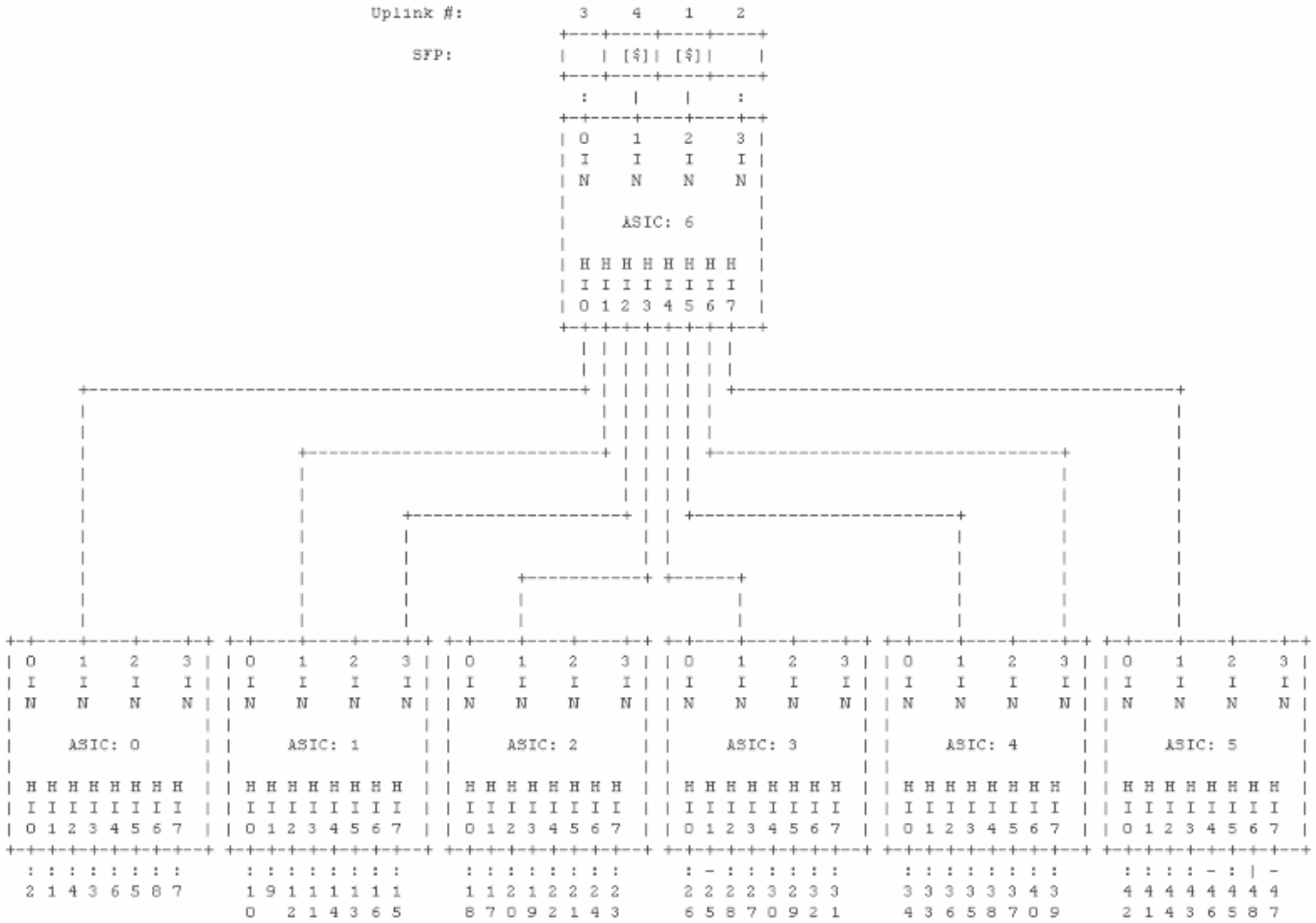
인터페이스 카운터 출력을 인터페트 하려면 전면 포트 번호를 HI 번호로 변환해야 합니다. 변환은 FEX 샤페 모델에 따라 달라집니다.

### N2K-C2148T-1GE

이 예에서 전면 포트 26(chassis-id/1/26)은 `rmon 3 HI 0`에 할당되었습니다.

```
switch# attach fex chassis_id
```

```
fex-[chassis_id]# show platform software redwood sts
```



## N2K-C2224TP-1GE / N2K-C2248TP-1GE

이 예에서는 전면 포트 10(135/1/10)에 HI 9가 할당되었습니다.

### 스위치 번호 연결 fex chassis\_id

fex-[chassis\_id]# dbgexec portola

### 포트> fp

```

fex-135# dbgexec prt
prt> fp
Fabric port map:
Fabric port map:
  1    3
  |    :
+-----+-----+
| NI1 | NIO |
+-----+-----+
| NI2 | NI3 |
+-----+-----+
  :    |
  2    4
Front port map:
  1  3  5  7  9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47
  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
+---+---+---+---+ +---+---+---+---+ +---+---+---+---+ +---+---+---+---+
HIF | 3 | 7 | 2 | 6 | 11 | 16 | | 10 | 15 | 17 | 20 | 21 | 23 | | 26 | 30 | 27 | 31 | 35 | 39 | | 34 | 38 | 42 | 46 | 43 | 47 |
+---+---+---+---+ +---+---+---+---+ +---+---+---+---+ +---+---+---+---+
HIF | 1 | 5 | 0 | 4 | 9 | 13 | | 8 | 12 | 14 | 18 | 19 | 22 | | 24 | 28 | 25 | 29 | 32 | 37 | | 33 | 36 | 40 | 44 | 41 | 45 |
+---+---+---+---+ +---+---+---+---+ +---+---+---+---+ +---+---+---+---+
  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
  2  4  6  8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48
prt> █

```

## N2K-C2232PP-10GE / N2K-C2232TM-10GE

이 예에서는 전면 포트 20(135/1/20)에 HI 19가 할당되었습니다.

```
switch# attach fex chassis_id
```

```
fex-[chassis_id]# show platform software woodside sts
```

```
(FINAL POSITION TBD)      Uplink #:      1  2  3  4  5  6  7  8
Link status:             :  :  :  :  :  :  :  |  |
SFP:                     [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]
                          +-----+-----+-----+-----+
                          | N  N  N  N  N  N  N  N  N  |
                          | I  I  I  I  I  I  I  I  I  |
                          | O  1  2  3  4  5  6  7  |
                          |                               |
                          |           NI (0-7)           |
                          +-----+-----+-----+
                          |                               |
                          |-----+-----+-----+-----+
                          |                               |
+-----+-----+-----+-----+-----+-----+-----+
|           HI (0-7)           | |           HI (8-15)           | |           HI (16-23)           | |           HI (24-31)           |
| H  H  H  H  H  H  H  H  | | H  H  H  H  H  H  H  H  | | H  H  H  H  H  H  H  H  | | H  H  H  H  H  H  H  H  |
| I  I  I  I  I  I  I  I  | | I  I  I  I  I  I  I  I  | | I  I  I  I  I  I  I  I  | | I  I  I  I  I  I  I  I  |
| O  1  2  3  4  5  6  7  | | 8  9  1  1  1  1  1  1  | | 1  1  1  1  2  2  2  2  | | 2  2  2  2  2  2  3  3  |
|                               | |           0  1  2  3  4  5  | | 6  7  8  9  0  1  2  3  | | 4  5  6  7  8  9  0  1  |
+-----+-----+-----+-----+-----+-----+-----+
[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]
+-----+-----+-----+-----+-----+-----+-----+
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
  1  2  3  4  5  6  7  8      9  1  1  1  1  1  1  1      1  1  1  2  2  2  2  2      2  2  2  2  2  3  3  3
                                0  1  2  3  4  5  6      7  8  9  0  1  2  3  4      5  6  7  8  9  0  1  2
```

## N2K-C2248TP-E-1G

```
fex-111# dbgexec pri
```

```
pri> fp
```

```
Fabric port map:
```

```
Fabric port map:
```

```
  1      3
  |      :
  +-----+-----+
  | NI1 | NI0 |
  +-----+-----+
  | NI2 | NI3 |
  +-----+-----+
  |      :
  2      4
```

```
Front port map:
```

```
  1  3  5  7  9  11  13  15  17  19  21  23  25  27  29  31  33  35  37  39  41  43  45  47
  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
HIF | 3 | 7 | 2 | 6 | 11 | 16 | 10 | 15 | 17 | 20 | 21 | 23 | 26 | 30 | 27 | 31 | 35 | 39 | 34 | 38 | 42 | 46 | 43 | 47 |
HIF | 1 | 5 | 0 | 4 | 9 | 13 | 8 | 12 | 14 | 18 | 19 | 22 | 24 | 28 | 25 | 29 | 32 | 37 | 33 | 36 | 40 | 44 | 41 | 45 |
  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
  2  4  6  8  10  12  14  16  18  20  22  24  26  28  30  32  34  36  38  40  42  44  46  48
```

## N2K-C2248PQ-10GE 및 N2K-C2348UPQ-10GE

이 예에서는 HI28이 전면 포트 29에 매핑됩니다.

tib> fp

-----			NI0,1	NI4,5
1 3 5 7 9 1 1 1   1 1 2 2 2 2 2 3   3 3 3 3 4 4 4 4	1-4	9-12		
1 3 5   7 9 1 3 5 7 9 1   3 5 7 9 1 3 5 7				
-----			-----	
H	H	H		
I	I	I		
0 2 4 6 8 1 1 1	1 1 2 2 2 2 2 3	3 3 3 3 4 4 4 4		
0 2 4	6 8 0 2 4 6 8 0	2 4 6 8 0 2 4 6		
-----			-----	
H	H	H		
I	I	I		
1 3 5 7 9 1 1 1	1 1 2 2 2 2 2 3	3 3 3 3 4 4 4 4		
1 3 5	7 9 1 3 5 7 9 1	3 5 7 9 1 3 5 7		
-----			-----	
2 4 6 8 1 1 1 1   1 2 2 2 2 2 3 3   3 3 3 4 4 4 4 4	5-8	13-16		
0 2 4 6	8 0 2 4 6 8 0 2	4 6 8 0 2 4 6 8		
-----			-----	
-----			NI2,3	NI6,7

## SFP 확인

이 명령은 포트에 대한 SFP(Small Form-Factor Pluggable) 정보를 표시합니다.

**fex# show platform software woodside sfp rmon 0 HI5**

이 예에서는 HI5의 SFP가 CISCO-AVAGO에서 만든 10G-Base-SR(LC)임을 확인할 수 있습니다.

```

## SFP Info:
  SFP FP-Port      : 0
  Fcot Num        : 0
  Fcot Type       : Not Found
10G-Base-SR      : Yes (Byte 3)
SONET            : No  (Bytes 4-5)
Ethernet         : No  (Byte 6)
FC               : No  (Bytes 7-10)
  SFP Type        : Gb Eth
  Min/Max Speeds  : [4294967295, 4294967295] Mbps

>> BASE ID FIELDS <<
Bytes  Name                Value
-----  ----
0       Identifier          : 0x03 (SFP Transceiver)
1       Ext. Identifier     : 0x04
2       Connector Type     : 0x07 (LC)
3-10    Transceiver        : 0x10 0x00 0x00 0x00 0x00 0x00 0x00 0x00
(4-5)   - SONET ComplCode  : 0x00 0x00 (None)
(6)     - Eth ComplCode   : 0x00 (Reserved)
(7)     - FC LinkLength   : 0x00 (None)
(7-8)   - FC TxType       : 0xFF (None)
(9)     - FC TxMedia      : 0x00 (None)
(10)    - FC Speed        : 0x00 (None)
11      Encoding          : 0x06 (64B/66B)
12      BR, Nominal       : 0x67
13      Reserved         : 0x00
14      Length(9m)-km    : 0x00
15      Length(9m)       : 0x00
16      Length(50m)      : 0x08
17      Length(62.5)     : 0x02
18      Length(Copper)   : 0x00
19      Reserved         : 0x1E
20-35   Vendor Name       : CISCO-AVAGO
36      Reserved         : 0x00
37-39   Vendor OUI       : 0x00 0x17 0x6A (0)
40-55   Vendor PN        : SFBR-7700SDZ
56-59   Vendor Rev       : 0x42 0x34 0x20 0x20 (B4 )
60-62   Reserved         : 0x03 0x52 0x00
63      CC_BASE          : 0x84

```

**참고:** 구리 포트를 사용하는 FEX에서 이 명령을 실행하면 명령 오류가 표시됩니다. 쿼리할 SFP가 없으므로 이 메시지가 표시됩니다. 해당 포트가 파이버이지만 현재 SFP가 없는 경우 SFP를 찾을 수 없습니다.

## 손실 찾기

FEX 패브릭 포트 링크의 FEX 쪽에 있는 인터페이스 카운터를 보려면 HI 및 NI 포트에 대한 FEX 프롬프트에서 show 명령을 실행할 수 있습니다.

### HI 포트 카운터 보기

이 명령은 show int와 유사한 포트 카운터 확인을 표시합니다.

```
fex-128# show platform software woodside rmon 0 HI0
```

TX	Current	Diff	Current	Diff	RX
TX_PKT_LT64	0	0	0	0	RX_PKT_LT64
TX_PKT_64	0	0	0	0	RX_PKT_64
TX_PKT_65	0	0	0	0	RX_PKT_65
TX_PKT_128	0	0	0	0	RX_PKT_128
TX_PKT_256	0	0	0	0	RX_PKT_256

**참고:** rmon 0은 FEX에 하나의 호스트 ASIC이 있는 경우에만 사용됩니다. 2224, 2248 및 2232 모델에는 하나의 기본 모델만 있습니다. 2148 모델에는 6개의 염소가 있으므로 rmon 0~5가 사용됩니다. 자세한 내용은 전면 포트 매핑 섹션을 참조하십시오.

## NI 포트 카운터 보기

이 명령은 **show int**와 유사한 네트워크 업링크에 대한 포트 카운터를 **표시합니다**. 이 명령은 링크의 FEX를 보여줍니다. 이 명령은 링크의 상위 스위치 측면을 표시하지 않습니다.

```
fex-128# show platform software woodside rmon 0 NI0
```

TX	Current	Diff	Current	Diff	RX
TX_PKT_LT64	0	0	0	0	RX_PKT_LT64
TX_PKT_64	0	0	0	0	RX_PKT_64
TX_PKT_65	0	0	0	0	RX_PKT_65
TX_PKT_128	0	0	0	0	RX_PKT_128
TX_PKT_256	0	0	0	0	RX_PKT_256

## 기록 삭제 보기

Historical Drops(기록 삭제)는 drops(삭제) 명령으로 볼 수 **있습니다**. FEX가 켜진 이후 모두 FEX에 드롭됩니다.

이 명령은 DROP8 카운터로 데이터 트래픽 삭제를 나타내지 않는 FEX CPU로 드롭하는 것도 보여줍니다. 이러한 작업은 무시해도 됩니다.

**참고:** tail drop [8] 및 TAIL\_DROP8은 FEX CPU에 대한 테일 드랍을 나타내며, 정상적인 조건



에서 발생할 경우 성능 문제를 해결할 수 없습니다.

```
prt> drops
PRT_SS_CNT_TAIL_DROP1 : 3 SS0
PRT_SS_CNT_TAIL_DROP1 : 6 SS1
PRT_SS_CNT_TAIL_DROP1 : 1 SS2
PRT_SS_CNT_TAIL_DROP1 : 25 SS3
PRT_SS_CNT_TAIL_DROP1 : 2 SS5
PRT_SS_CNT_TAIL_DROP8 : 142 SS0
PRT_SS_CNT_TAIL_DROP8 : 73 SS1
PRT_SS_CNT_TAIL_DROP8 : 11 SS2
PRT_SS_CNT_TAIL_DROP8 : 62048 SS3
PRT_SS_CNT_TAIL_DROP8 : 4613 SS4
PRT_SS_CNT_TAIL_DROP8 : 552 SS5
```

## 최근 삭제 및 인터럽트 보기

CPU로 전송된 인터럽트에는 정체 및 버퍼 공간 부족으로 인해 삭제된 tail-drop이 포함됩니다. `show new_ints` 명령으로 볼 수 있습니다.

**참고:** 6.0 이상 코드에서는 `show new_ints all`을 사용합니다.

다음 예에서는 SS1 버퍼에서 프레임 테일이 드롭되는 것을 보여줍니다.

```
prt> show new_ints
|-----|
| SS1 : ssx_int_norm_td
|-----+-----+
| 1 | 00001c98 | tail drop[1] | frames are being tail dropped.
| 2 | 00005cac | tail drop[2] | frames are being tail dropped.
| 8 | 0000012e | tail drop[8] | frames are being tail dropped.
```

다음 예에서는 NI 3에서 기호 오류를 수신함을 보여 줍니다.

```
| NI3 : nix_xe_INT_xg
|-----+-----+
|2 |00000005 | rx_local_fault | Link is in local fault state
|3 |00000007 | rx_remote_fault | Link is in remote fault state
|4 |00000004 | rx_code_violation | MAC received unexpected XGMII control characters.
|5 |00000004 | rx_err_symbol | MAC received an XGMII error character.
|16|00000001 | rx_local_fault_edge | Local fault state has changed.
|17|00000001 | rx_remote_fault_edge | Remote fault state has changed.
|-----|
```

다음 예에서는 FEX 테일이 인그레스 NI3를 삭제하는 프레임을 삭제하는 것을 보여 줍니다.

```
| SS4 : ssx_int_err
|-----+-----+
|0 |00031aa9 | wo_cr[0] | frames rcvd without credit for pausable classes. Pause
is missing.
|1 |00014e21 | wo_cr[1] | frames rcvd without credit for pausable classes. Pause
is missing.
|2 |00018a9f | wo_cr[2] | frames rcvd without credit for pausable classes. Pause
is missing.
|3 |00025efb | wo_cr[3] | frames rcvd without credit for pausable classes. Pause
is missing.
```

|-----|

## 실시간으로 포트 트래픽 속도 보기

rate 명령은 포트에 대한 실시간 트래픽 속도 통계를 출력합니다. **show int**와 달리 평균이 아닌 원시 현재 데이터 속도입니다. 이 예에서 NI 3은 현재 네트워크 대 호스트 방향에서 2.96kbps를 수신합니다. 해당 상위 Nexus 스위치의 **show int**는 NI 3에 연결된 FEX 패브릭 업링크의 TX 방향에서 2.96Kbps를 표시합니다.

```
prt> rate
```

Port	Tx Packets	Tx Rate (pkts/s)	Tx Bit Rate	Rx Packets	Rx Rate (pkts/s)	Rx Bit Rate	Avg Pkt (Tx)	Avg Pkt (Rx)	Err
O-CI	11	2	4.80Kbps	12	2	8.64Kbps	252	430	
O-NI3	6	1	4.32Kbps	6	1	2.96Kbps	430	289	
O-NI1	6	1	4.32Kbps	5	1	1.89Kbps	430	217	

## 손실 완화

테일 드롭은 버퍼 소모 때문에 발생합니다. 일반적으로 여러 서버가 동시에 HIF로 버스트되거나 호스트 이그레스 버퍼가 아웃바운드 트래픽을 빠르게 비울 수 없어 NIF의 크레딧을 보충할 수 있습니다.

이러한 손실을 완화하는 몇 가지 옵션이 있습니다.

## 서버 위치 변경

스토리지 어레이 및 비디오 엔드포인트와 같은 버스트 트래픽 폴로우가 있는 서버를 FEX에서 분리한 다음 상위 스위치의 기본 포트에 직접 연결합니다. 이렇게 하면 버스트 서버가 버퍼를 소모하고 덜 소란스러운 호스트에서 트래픽을 소진하는 것을 방지할 수 있습니다.

Nexus 5000 및 6000 Series 스위치는 FEX 모델보다 더 큰 버퍼를 가지고 있어, 버스트 서버를 기본 포트에 연결하여 더 큰 버스트를 처리할 수 있기 때문에 손실을 완화합니다.

## 추가 업링크 추가

일부 FEX 모델은 FEX에서 상위 스위치로의 업링크가 더 추가될 때 추가 버퍼 공간을 잠금 해제할 수 있습니다. 이는 네트워크 업링크에서 삭제될 수 있습니다.

### 표 2.

모델	업링크를 추가할 때 버퍼 증가
2148	없음
2224	최대 2개의 업링크
2248TP	최대 4개의 업링크
2232	최대 4개의 업링크
2248TP-E	없음
2248PQ	없음

## HI 버퍼 공유

대부분의 FEX 모델은 모든 호스트 포트에서 HI 버퍼를 공유함으로써 이점을 얻을 수 있습니다. HI에서 드롭이 확인되면 버퍼를 공유하면 이러한 드롭이 완화될 수 있습니다.

**FEX 대기열 제한을 전역적으로 수정합니다.**

5k(config) # fex queue-limit 없음(해당 5k의 모든 fex에 전역적으로 적용)

**개별 FEX의 FEX 큐 제한을 수정합니다.**

Fex 큐

5k(config) # fex 100

5k(config-fex)# 하드웨어 [model] queue-limit 없음

## Nexus 6000 FEX 로드 밸런스 향상

Nexus 6000에는 로드 밸런싱 알고리즘을 HIF에서 NIF로 변경하는 추가 옵션이 있습니다. 기본적으로 패킷이 서로 다른 HIF 포트에 도착하더라도 동일한 NIF로 대기될 수 있습니다. 업링크 로드 밸런싱 모드가 활성화되면 여러 NIF에 분산되며 NIF 이그레스 버퍼의 사용량도 늘릴 수 있습니다.

6k(config)# 하드웨어 N2248PQ 업링크 로드 밸런싱 모드