

排除Nexus 5000/6000系列上的FEX效能問題

目錄

[簡介](#)

[背景資訊](#)

[導航CLI](#)

[連線到FEX](#)

[進入調試執行模式](#)

[退出調試執行模式](#)

[退出FEX](#)

[技術](#)

[主機介面\(HI\)](#)

[網路介面\(NI\)](#)

[FEX光纖連線埠](#)

[FEX ASIC名稱](#)

[前埠對映](#)

[N2K-C2148T-1GE](#)

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

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

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

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

[驗證SFP](#)

[查詢損失](#)

[檢視HI埠計數器](#)

[檢視NI埠計數器](#)

[檢視歷史丟棄](#)

[檢視最近的丟棄和中斷](#)

[即時檢視埠流量速率](#)

[減少損失](#)

[重新定位伺服器](#)

[新增其他上行鏈路](#)

[共用HI緩衝區](#)

[Nexus 6000 FEX負載平衡增強功能](#)

簡介

本文說明如何對可連線到Nexus 5000或6000系列交換機的交換矩陣擴展器(FEX)的效能進行故障排除。

注意：本文檔中引入的命令均不具有中斷性。 必須將Nexus 2000交換機連線到5000或6000系列交換機。

背景資訊

導航CLI

連線到FEX

連線到FEX以在FEX命令列上運行show命令：

```
Nexus#連線fex fex  
fex>
```

進入調試執行模式

在FEX上進入debug模式以運行高級命令並指定FEX基本名稱。有關FEX基本名稱，請參閱表1。

```
fex# dbgexec [prt/woo/red/pri]
```

退出調試執行模式

要退出調試執行模式，請使用CTRL+C鍵盤序列：

```
fex> [CTRL+C]
```

退出FEX

要退出fex，請使用命令exit:

```
fex#退出
```

技術

主機介面(HI)

Hi是FEX上面向伺服器的埠。這些埠通常稱為前端埠。FEX上的每個前端埠都有一個高可用性編號。此編號通常與埠編號不同，但用於對引用埠的命令進行故障排除。每個asic以不同的方式將前埠清單顯示。

網路介面(NI)

NI是連線回父交換機的FEX控制埠。這些也稱為網路上行鏈路。它們還具有取決於模型的唯一NI編號。

FEX光纖連線埠

這些埠是FEX唯一鏈路的父交換機端。這些埠使用switchport mode fex-fabric和fex association命令進行配置。

FEX ASIC名稱

每個FEX設計有不同的ASIC。ASIC名稱的縮寫在調試模式下用於運行命令。

大多數FEX型號都有一個ASIC，但2148有6個，每個有8個前埠。這些在故障排除命令中稱為rmon。

列出了ASIC名稱及相關縮寫以供參考：

表1.

FEX型號	ASIC名稱	縮寫
N2K-C2148T-1GE	紅杉木	rw
N2K-C2224TP-1GE	波托拉	prt
N2K-C2248TP-1GE		
N2K-C2232PP-10GE	伍德賽德	woo
N2K-C2232TM-10GE		
N2K-C2248TP-E-1GE	普林斯頓	pri
B22	伍德賽德	woo
N2K-C2232TM-E-10GE	伍德賽德	woo
N2K-C2248PQ-10GE	伍德賽德/貝爾蒙特	woo
N2K-C2348UPQ-10GE	蒂布龍	tib

前埠對映

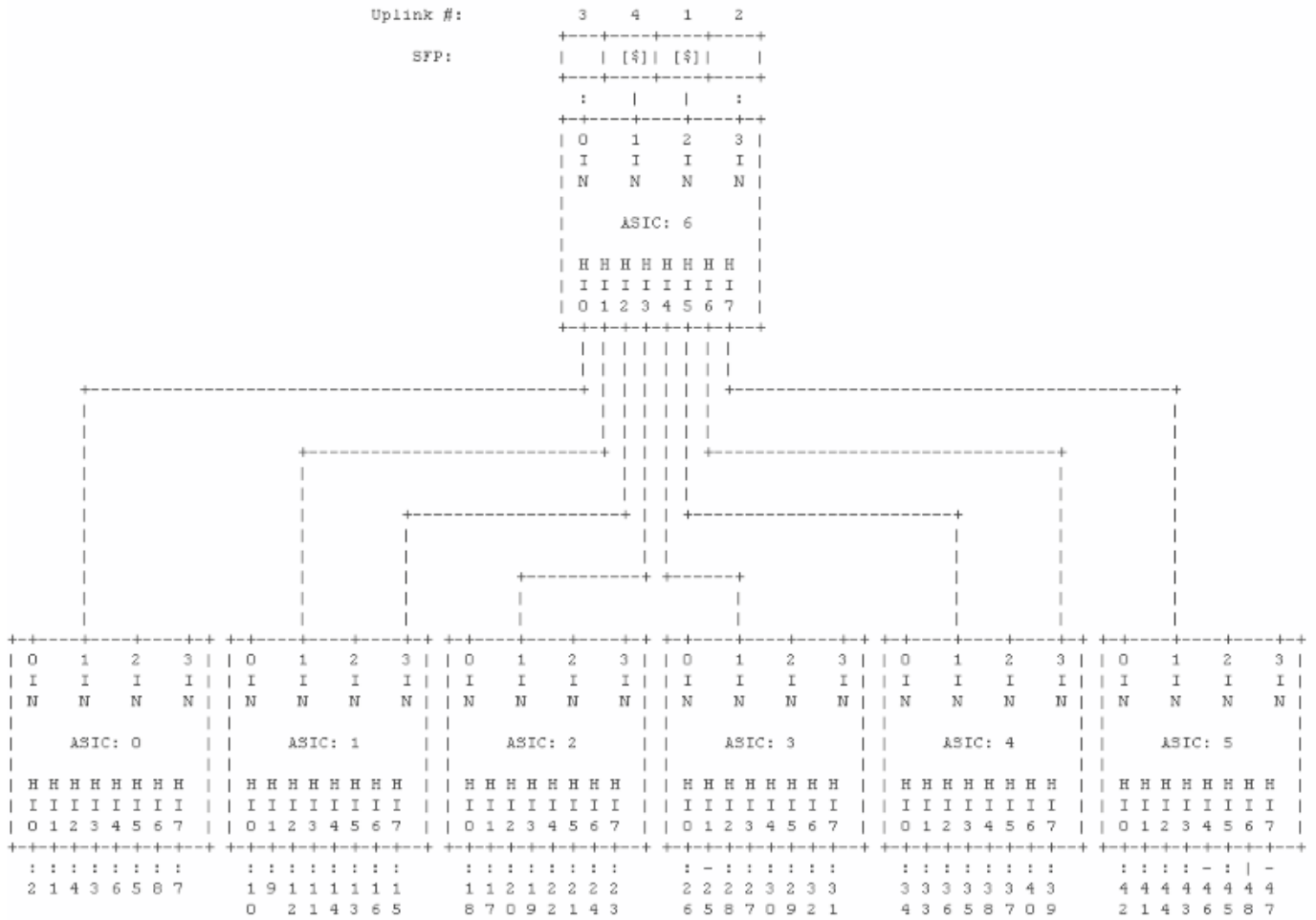
要插入介面計數器輸出，可能需要將前埠號轉換為HI號。轉換取決於FEX機箱型號。

N2K-C2148T-1GE

在本示例中，前埠26 (機箱ID/1/26) 已分配為3 HI 0:

```
switch# attach fex chassis_id
```

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



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

在本例中，為前埠10(135/1/10)分配了HI 9:

```
switch# attach fex chassis_id
```

```
fex-[機箱_id]# dbgexec portola
```

```
prt> 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

在本示例中，前埠20(135/1/20)已分配為HI 19:

```
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 |
                         | 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  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和N2K-C2348UPQ-10GE

在本例中，HI28對映到前埠29:

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	+-----+	+-----+
	0 2 4 6	8 0 2 4 6 8 0 2	5-8	13-16
		4 6 8 0 2 4 6 8		
+-----+-----+-----+			+-----+	+-----+
-----			NI2,3	NI6,7

驗證SFP

此命令顯示埠的小型封裝熱插拔(SFP)資訊。

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

在此範例中，您會看到HI5中的SFP是由CISCO-AVAGO製造的10G-Base-SR(LC):

```

## 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

```

附註： 如果在使用銅纜埠的FEX上運行此命令，則會發現命令錯誤。這是預期結果，因為沒有要查詢的SFP。當該埠是光纖但當前不包含SFP時，該提示將返回**no SFP found**。

查詢損失

可以在FEX提示符下為HI和NI埠運行show命令，以便檢視FEX交換矩陣埠鏈路FEX端的介面計數器。

檢視HI埠計數器

此命令顯示埠計數器驗證，類似於show int:

```
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

附註：僅當FEX具有一個主機asic時才使用rmon 0。2224、2248和2232型號只有一個asic。2148型號有6個asic，因此使用rmon 0至5。請參見「前埠對映」部分以瞭解更多詳細資訊。

檢視NI埠計數器

此命令會顯示類似show int的網路上行鏈路的埠計數器。此命令會顯示鏈路的FEX端。此命令不會顯示鏈路的父交換機端。

```
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

檢視歷史丟棄

可以使用drops命令檢視歷史丟包。這將顯示FEX開啟後所有丟棄的情況。

此命令還顯示丟棄到FEX CPU，而不表示使用DROP8計數器丟棄的資料流量。可以放心地忽略這些。

註:tail drop [8]和TAIL_DROP8表示FEX CPU的尾部丟棄，與排除效能無關，因為正常情況下會發生這種情況。


```
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
```

檢視最近的丟棄和中斷

傳送到CPU的中斷包括尾部丟棄，尾部丟棄是由於擁塞和缺少緩衝空間而丟棄的。可以使用**show new_ints**命令檢視這些專案：

注意:6.0及更高版本的代碼使用**show new_ints all**

此範例顯示SS1緩衝區中的訊框尾部捨棄：

```
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.
```

此示例顯示NI 3收到符號錯誤：

```
| 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.
|-----|
```

此範例顯示FEX尾部捨棄輸入NI3的訊框：

```
| 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.
|-----|
```

即時檢視埠流量速率

rate命令輸出埠的即時流量速率統計資訊。與show int不同，它不是平均值，而是第二位的原始當前資料速率。在本示例中，NI 3當前在從網路到主機的方向上接收2.96kbps。在連線到NI 3的FEX交換矩陣上行鏈路上，相應父Nexus交換機上的show int在TX方向上顯示2.96Kbps。

```
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	

減少損失

尾部丟棄是由緩衝區耗盡引起的。通常，當多個伺服器一次突發到HIF，或者主機出口緩衝區無法以足夠快的速度清空出站流量來補充NIF上的信用時，緩衝區會耗盡。

有幾種辦法可以減輕這一損失。

重新定位伺服器

將具有突發資料流的所有伺服器（例如儲存陣列和影片終端）移出FEX，並將其直接連線到父交換機的基本埠。這樣可以防止突發伺服器耗盡緩衝區，並耗盡來自非活躍主機的流量。

Nexus 5000和6000系列交換機的緩衝區比FEX型號大，可將突發伺服器連線到基本埠，從而減少損失，因為基本埠緩衝區能夠處理更大規模的突發流量。

新增其他上行鏈路

當從FEX到父交換機的更多上行鏈路被新增時，某些FEX型號可以釋放額外的緩衝區空間。這可能會停止網路上行鏈路上的捨棄專案。

表2.

型號	新增上行鏈路時緩衝區增加
2148	none
2224	緩衝區增加多達2個上行鏈路
2248TP	緩衝區增加多達4個上行鏈路
2232	緩衝區增加多達4個上行鏈路
2248TP-E	none
2248PQ	none

共用HI緩衝區

大多數型號的FEX都可以通過在所有主機埠上共用高可用性緩衝區而受益。如果在HI上看到丟包，共用緩衝區可能會減輕這些丟包。

全域性修改FEX隊列限制：

5k(config)# no fex queue-limit (全域性適用於該5k上的所有副本)

修改單個FEX上的FEX隊列限制：

Fex隊列

5k(config)# fex 100

5k(config-fex)# no hardware [model] queue-limit

Nexus 6000 FEX負載平衡增強功能

Nexus 6000具有將負載平衡演算法從HIF更改為NIF的附加選項。預設情況下，即使資料包到達不同的HIF埠，它們仍可能排隊到相同的NIF。啟用上行鏈路負載均衡模式後，它們分佈在多個NIF上，並允許更均勻地使用NIF輸出緩衝區。

6k(config)# hardware N2248PQ uplink-load-balance-mode