

Solucione problemas de desempenho do FEX no Nexus 5000/6000 Series

Contents

[Introduction](#)

[Informações de Apoio](#)

[Navegar pelo CLI](#)

[Anexar ao FEX](#)

[Entre no modo exec de depuração](#)

[Sair do modo exec de depuração](#)

[Sair do FEX](#)

[Terminology](#)

[Interface de host \(HI\)](#)

[Interface de rede \(NI\)](#)

[Porta de estrutura FEX](#)

[Nomes ASIC FEX](#)

[Mapeamento de porta frontal](#)

[N2K-C2148T-1GE](#)

[N2K-C224TP-1GE / N2K-C2248TP-1GE](#)

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

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

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

[Verificar SFP](#)

[Localizar perda](#)

[Exibir contadores de porta HI](#)

[Exibir contadores de porta NI](#)

[Exibir quedas históricas](#)

[Exibir quedas e interrupções recentes](#)

[Exibir a taxa de tráfego de porta em tempo real](#)

[Reduza a perda](#)

[Reposicionar servidores](#)

[Adicionar uplinks adicionais](#)

[Compartilhar buffers HI](#)

[Aprimoramento do balanceamento de carga do Nexus 6000 FEX](#)

Introduction

Este documento descreve como solucionar problemas de desempenho nos extensores de estrutura (FEX) que podem ser anexados aos switches Nexus 5000 ou 6000 Series.

Observação: nenhum dos comandos apresentados neste documento apresenta interrupções. Você deve ter um switch Nexus 2000 conectado a um switch 5000 ou 6000

Series.

Informações de Apoio

Navegar pelo CLI

Anexar ao FEX

Conecte-se ao FEX para executar os comandos show na linha de comando do FEX:

```
Nexus#attach fex fex  
fex>
```

Entre no modo exec de depuração

Entre no modo de depuração no FEX para executar comandos avançados e especificar o nome básico do FEX. Consulte a Tabela 1. para os nomes básicos do FEX.

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

Sair do modo exec de depuração

Para sair do modo exec de depuração, use a sequência de teclado CTRL+C:

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

Sair do FEX

Para sair do fex, use o comando **exit**:

```
fex# exit
```

Terminology

Interface de host (HI)

Olá são as portas que enfrentam os servidores no FEX. Elas são comumente conhecidas como portas frontais. Cada porta frontal em um FEX tem um número HI. Esse número geralmente é diferente do número da porta, mas é usado para solucionar problemas de comandos para se referir a uma porta. Cada básico tabula as portas frontais de forma diferente.

Interface de rede (NI)

NIs são as portas de controle FEX no FEX que se conectam de volta ao switch pai. Eles também são conhecidos como uplinks de rede. Eles também têm um número de NI exclusivo dependente

do modelo.

Porta de estrutura FEX

Essas portas são o lado do switch pai do link exclusivo para o FEX. Essas portas são configuradas com os comandos **switchport mode fex-fabric** e **fex association**.

Nomes ASIC FEX

Cada FEX é projetado com um ASIC diferente. A abreviação do nome ASIC é usada no modo de depuração para executar comandos.

A maioria dos modelos do FEX tem um ASIC, mas o 2148 tem 6, cada um com 8 portas frontais. Eles são chamados de **rmon** em comandos de identificação e solução de problemas.

Os nomes ASIC e abreviações associadas estão listados para referência:

Tabela 1.

Modelo FEX	Nome ASIC	Abreviação
N2K-C2148T-1GE	sequoia	rw
N2K-C224TP-1GE	portola	porta
N2K-C2248TP-1GE		
N2K-C2232PP-10GE	madeira	woo
N2K-C2232TM-10GE		
N2K-C2248TP-E-1GE	princeton	pri
B22	madeira	woo
N2K-C2232TM-E-10GE	madeira	woo
N2K-C2248PQ-10GE	floresta/belmont	woo
N2K-C2348UPQ-10GE	tiburão	ponta

Mapeamento de porta frontal

Para converter o número da porta frontal em um número HI, pode ser necessário fazer a conversão da saída do contador de interface de intervalo. A conversão depende do modelo de chassi FEX.

N2K-C2148T-1GE

Neste exemplo, a porta frontal 26 (chassis-id/1/26) recebeu rmon 3 HI 0:

```
switch# attach fex chassis_id
```

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

```

Uplink #:          3    4    1    2
                +---+---+---+---+
                | | [§] | [§] | |
                +---+---+---+---+
                : | | :
                +---+---+---+---+
                | 0  1  2  3 |
                | I  I  I  I |
                | N  N  N  N |
                |
                | ASIC: 6
                |
                | H H H H H H H H |
                | I I I I I I I I |
                | 0 1 2 3 4 5 6 7 |
                +---+---+---+---+
                | | | | | | | |
                | | | | | | | |
                | | | | | | | |
                | | | | | | | |
                | | | | | | | |
                | | | | | | | |
                +---+---+---+---+
                | 0  1  2  3 | | 0  1  2  3 | | 0  1  2  3 | | 0  1  2  3 | | 0  1  2  3 | | 0  1  2  3 |
                | I  I  I  I | | I  I  I  I | | I  I  I  I | | I  I  I  I | | I  I  I  I | | I  I  I  I |
                | N  N  N  N | | N  N  N  N | | N  N  N  N | | N  N  N  N | | N  N  N  N | | N  N  N  N |
                |
                | ASIC: 0 | | ASIC: 1 | | ASIC: 2 | | ASIC: 3 | | ASIC: 4 | | ASIC: 5 |
                |
                | 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 | | 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 | | I I I I I I I I | | I I I I I I I I |
                | 0 1 2 3 4 5 6 7 | | 0 1 2 3 4 5 6 7 | | 0 1 2 3 4 5 6 7 | | 0 1 2 3 4 5 6 7 | | 0 1 2 3 4 5 6 7 | | 0 1 2 3 4 5 6 7 |
                +---+---+---+---+ +---+---+---+---+ +---+---+---+---+ +---+---+---+---+ +---+---+---+---+ +---+---+---+---+
                : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
                2 1 4 3 6 5 8 7   1 9 1 1 1 1 1 1   1 1 2 1 2 2 2 2   2 2 2 2 3 2 3 3   3 3 3 3 3 3 4 3   4 4 4 4 4 4 4 4
                0 2 1 4 3 6 5   8 7 0 9 2 1 4 3   6 5 8 7 0 9 2 1   4 3 6 5 8 7 0 9   2 1 4 3 6 5 8 7

```

N2K-C224TP-1GE / N2K-C2248TP-1GE

Neste exemplo, a porta frontal 10 (135/1/10) foi atribuída à HI 9:

```
switch#attach fex chassis_id
```

```
fex-[chassis_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

Neste exemplo, a porta frontal 20 (135/1/20) recebeu a 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  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 e N2K-C2348UPQ-10GE

Neste exemplo, o HI28 mapeia para a porta frontal 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	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

Verificar SFP

Esse comando mostra as informações SFP (Small Form-Factor Pluggable) da porta.

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

Neste exemplo, você vê que o SFP em HI5 é um 10G-Base-SR (LC) feito pela CISCO-AVAGO:

```

## 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

```

Observação: se você executar esse comando em um FEX que usa portas de cobre, você observará os erros de comando. Isso é esperado porque não há SFP para consulta. O prompt retornará para **nenhum SFP encontrado** quando essa porta for de fibra, mas não contiver atualmente um SFP.

Localizar perda

Os comandos show podem ser executados no prompt FEX para as portas HI e NI para visualizar contadores de interface no lado FEX dos links de porta FEX Fabric.

Exibir contadores de porta HI

Nota: queda traseira [8] e TAIL_DROP8 representam quedas traseiras para a CPU FEX e não são relevantes para a identificação e solução de problemas de desempenho, pois isso acontece em condições normais.

```
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
```

Exibir quedas e interrupções recentes

As interrupções enviadas para a CPU incluem quedas traseiras, que são quedas devido a congestionamento e falta de espaço no buffer. Eles podem ser visualizados com o comando **show new_points**:

Nota: o código 6.0 e posterior usa **show new_points all**

Este exemplo mostra que a cauda dos quadros no buffer SS1:

```
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.
```

Este exemplo mostra que o NI 3 recebe erros de símbolo:

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

Este exemplo mostra que a cauda FEX descarta quadros que ingressam em NI3:

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

Exibir a taxa de tráfego de porta em tempo real

O comando `rate` gera estatísticas de taxa de tráfego em tempo real para uma porta. Ao contrário do `show int`, não é uma média, é a taxa de dados atual bruta que está em segundo lugar. Neste exemplo, o NI 3 recebe atualmente 2,96 kbps na direção Rede ao Host. Um `show int` no switch Nexus pai correspondente mostra 2,96 Kbps na direção TX no uplink de estrutura FEX conectado ao NI 3.

```
prt> rate
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Port  || Tx Packets | Tx Rate | Tx Bit  || Rx Packets | Rx Rate | Rx Bit  | Avg Pkt| Avg Pkt| |
|      ||            | (pkts/s) | Rate    ||            | (pkts/s) | Rate    | (Tx)   | (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    | |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
```

Reduza a perda

As quedas na cauda são causadas pelo esgotamento do buffer. Geralmente, o buffer se torna esgotado quando vários servidores são interrompidos para os HIFs de uma só vez, ou os buffers de saída do host não podem esvaziar seu tráfego de saída rápido o suficiente para repor os créditos nos NIFs.

Há várias opções disponíveis para reduzir essa perda.

Reposicionar servidores

Mova todos os servidores com fluxos de tráfego intermitentes, como storage arrays e endpoints de vídeo, do FEX e conecte-os diretamente às portas base do switch pai. Isso impedirá que os servidores em surtos esgotem o buffer e eliminem o tráfego de hosts com menos bate-papo.

Os switches Nexus 5000 e 6000 Series têm buffers maiores do que os modelos FEX, para conectar servidores em surtos às portas de base reduz a perda porque os buffers de porta de base são capazes de lidar com uma intermitência muito maior.

Adicionar uplinks adicionais

Alguns modelos de FEX podem liberar espaço de buffer adicional quando mais uplinks do FEX para o switch pai são adicionados. Isso pode potencialmente parar quedas nos uplinks de rede.

Tabela 2.

Modelo	Aumento do buffer ao adicionar uplinks
2148	nenhum

2224 aumento de buffer de até 2 uplinks
2248TP aumento de buffer de até 4 uplinks
2232 aumento de buffer de até 4 uplinks
2248TP-E nenhum
2248PQ nenhum

Compartilhar buffers HI

A maioria dos modelos de FEX pode se beneficiar do compartilhamento do buffer HI em todas as portas de host. Se houver descartes na interface de usuário (HI), o compartilhamento do buffer pode atenuar essas quedas.

Modificar o limite de fila FEX globalmente:

5k(config)# no fex queue-limit (aplica-se globalmente a todos os fexos nesse 5k)

Modificar o limite de fila FEX em FEX individual:

fila Fex

5k(config)#fex 100

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

Aprimoramento do balanceamento de carga do Nexus 6000 FEX

O Nexus 6000 tem uma opção adicional para alterar o algoritmo de balanceamento de carga de HIF para NIF. Por padrão, mesmo que os pacotes cheguem a portas HIF diferentes, eles ainda podem ser enfileirados para o mesmo NIF. Com o modo de balanceamento de carga de uplink ativado, eles são distribuídos por vários NIFs e permitem uma utilização mais uniforme dos buffers de saída NIF.

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