

Содержание

[Введение](#)

[Предварительные условия](#)

[Требования](#)

[Используемые компоненты](#)

[Входные буфера](#)

[Причины для паузы Тх на Nexus 2232](#)

[Случаи лабораторного испытания](#)

[Тест 1: Пульсирующий трафик с Управлением потоками, не включенным на хосте.](#)

[Тест 2: Пульсирующий трафик с Управлением потоками, включенным на хосте](#)

[Тест 3: хэш-коллизия EtherChannel](#)

[Заключения и оптимальные методы](#)

[Связанные обсуждения Сообщества Cisco Support](#)

Введение

Документ предоставит сведения, чтобы помочь устранять неполадки пауз ТХ на Nexus 2232 NIF (Интерфейс хоста) порты. Это фокусирует на трафике в H2N (Хост Сети) направление (трафик, входящий от серверов к сети, с юга на север). Это не покрывает сценарии, отнесенные к N2N (Сеть для Хостинга) трафики.

Этот документ прежде всего создан для Nexus 2232 FEX, но понятие просит B22 и 2248UPQ FEX (Модуль ввода-вывода).

Предварительные условия

Требования

Cisco рекомендует иметь базовые знания об этих темах

- Cisco Nexus конфигурация серии 2000
- Cisco Nexus конфигурация серии 6000

Используемые компоненты

Сведения, содержащиеся в данном документе, касаются следующих версий программного обеспечения и оборудования:

- Cisco Nexus N2K-C2232PP-10GE
- Cisco Nexus 6001
- 7.1 (1) N1 (1)

Сведения, представленные в этом документе, были получены от устройств, работающих в специальной лабораторной среде. Все устройства, описанные в этом документе, были запущены с чистой (стандартной) конфигурацией. В рабочей сети необходимо изучить потенциальное воздействие всех команд до их использования.

Входные буфера

Nexus 2232 имеет 32 1/10G NIF (направление хоста) порты и 8 10G NIF (Направление сети) порты.

Перед глубоким дайвингом в проблему паузы Tx у вас должно быть понимание доступных буферов на интерфейсах Fex. Буферы, назначенные на interface/qos-group, могут быть проверены через следующую команду на родительском коммутаторе

```
esc-6001# show queuing interface ethernet 147/1/1
```

```
if_slot 79, ifidx 0x1f920000
Ethernet147/1/1 queuing information:
Input buffer allocation:
Qos-group: 0
frh: 8
drop-type: drop
cos: 0 1 2 3 4 5 6
xon      xoff      buffer-size
-----+-----+-----
0        126720    151040
```

<snip>

Как замечено выше, с QoS (качествами обслуживания) по умолчанию, для трафика класса отбрасывания (группа qos 0), FEX NIF имеет 151040 байтов для буферизации трафика H2N, и порог XOFF составляет 126720 байтов.

Конфигурация flowcontrol

Nexus 2232 превышен 8:1. Для предотвращения отбрасывания пакета в направлении H2N из-за превышения подписки и переполнения буфера Nexus 2232 имеет flowcontrol NIF, пересылают по умолчанию.

```
esc-6001# show run int ethernet 147/1/1 all | inc flow
priority-flow-control mode auto
flowcontrol receive off
flowcontrol send on
esc-6001# show interface ethernet 147/1/1 flowcontrol
```

```
-----
Port          Send FlowControl Receive FlowControl RxPause TxPause
              admin   oper    admin   oper
-----
Eth147/1/1  on     on     off     off      0         0
```

Причины для паузы Tx на Nexus 2232

Когда порог XOFF 126720 байтов будет поражен, Nexus 2232 передаст Паузу Tx к хосту на NIF. Типичные причины для этого

1) Трафик H2N, входящий в FEX, является очень пульсирующим порождением входных буферов быть полным заполнением и ударом порога XOFF.

2) Большая часть FEX, развертывания используют port-channel для агрегации множественного NIFs. Пауза tx будет также замечена из-за входных буферов, получающих полную должную etherchannel хэш-коллизию на FEX. Это происходит, когда множественный, порты NIF пробуют к выходу из сингла NIF из-за результатов etherchannel.

Падения направления H2N

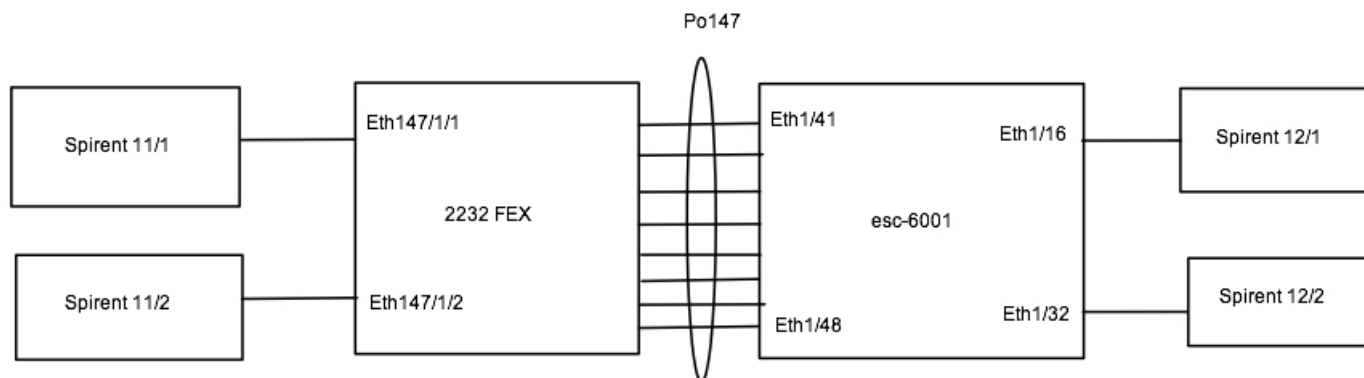
Для предотвращения отбрасывания пакета Пауза Tx передается, как только поражен порог XOFF. Однако отбрасывания трафика H2N могут быть замечены если

1) серверы не соблюдают паузу

2) или имейте задержку удаствивания паузы, заставляющей порог отбрасывания 151 КБ быть пораженным.

Случай лабораторного испытания

Топология



Для этого теста в лабораторной работе, существуют четыре 10G spirent порты, действующие как hosts, Два находятся на FEX, и два находятся на родительском Nexus 6001. Все порты находятся в VLAN 50. Нет никакого другого порта, активного на FEX или родителе.

```

esc-6001# show run int ethernet 147/1/1 all | inc flow
priority-flow-control mode auto
flowcontrol receive off
flowcontrol send on
esc-6001# show interface ethernet 147/1/1 flowcontrol
  
```

```

-----
Port          Send FlowControl Receive FlowControl RxPause TxPause
              admin  oper    admin  oper
-----
Eth147/1/1   on    on      off    off      0        0
  
```

Тест 1: Пульсирующий трафик с Управлением потоками, не включенным на хосте.

Передача 100K 1500-байтовая "скорость линии" одноадресно передала пакеты от хоста на Eth147/1/1 (к Eth1/16), и Eth147/1/9 (сделайте Eth1/32). Каждый поток является единым потоком. Управление потоками отключено на хосте (Spirent).

Результаты: порты Получателя сообщили о приблизительно 563 пакетах, отброшенных для каждого потока. Так как Управление потоками отключено на хосте, вы видите партию больше Паузы Tx и также большая задержка (приблизительно 100 микро секунд)

```
esc-6001# clear counters ; clear qos stat
esc-6001#
esc-6001# show interface ethernet 147/1/1, ethernet 147/1/9 | egrep Ethernet147|pause|unicast
Ethernet147/1/1 is up
  100000 unicast packets 0 multicast packets 0 broadcast packets
  0 Rx pause
  0 unicast packets 578269 multicast packets 0 broadcast packets
  578267 Tx pause
Ethernet147/1/9 is up
  100000 unicast packets 0 multicast packets 0 broadcast packets
  0 Rx pause
  0 unicast packets 578275 multicast packets 0 broadcast packets
  578273 Tx pause
esc-6001# show interface ethernet 147/1/1, eth147/1/9 flowcontrol
```

```
-----
Port          Send FlowControl Receive FlowControl RxPause  TxPause
          admin   oper    admin   oper
-----
Eth147/1/1  on     on     off     off     0         578267
Eth147/1/9  on     on     off     off     0         578273
-----
```

Отбрасывания, о которых сообщает получатель, фактически являются dropend на самом fex. Существуют внутренние hardware команды, которые могут показать отбрасывания, но это потребовало бы полного понимания fex внутренней архитектуры, которая выходит за рамки этого документа. Если необходимо проверить эти счетчики, затроньте TAC для этого аспекта устранения проблем.

Тест 2: Пульсирующий трафик с Управлением потоками, включенным на хосте

Передача единого потока 100К 1500-байтовая "скорость линии" одноадресно передала пакеты от хоста на Eth147/1/1 (к Eth1/16), и Eth147/1/9 (сделайте Eth1/32). Каждый поток является единым потоком. Управление потоками включено на хосте (Spirent).

Результаты: порты Получателя не сообщают ни о какой потере. Минимальная Пауза Tx и среднее число latency составляют приблизительно 19 микросекунд

```
esc-6001# clear counters ; clear qos stat
esc-6001# show interface ethernet 147/1/1, ethernet 147/1/9 | egrep Ethernet147|pause|unicast
Ethernet147/1/1 is up
  100000 unicast packets 0 multicast packets 0 broadcast packets
  0 Rx pause
  0 unicast packets 4743 multicast packets 0 broadcast packets
  4739 Tx pause
Ethernet147/1/9 is up
  100000 unicast packets 0 multicast packets 0 broadcast packets
  0 Rx pause
  0 unicast packets 4703 multicast packets 0 broadcast packets
  4700 Tx pause

esc-6001# show interface ethernet 147/1/1, eth147/1/9 flowcontrol
```

```
-----
Port          Send FlowControl Receive FlowControl RxPause  TxPause
          admin   oper    admin   oper
-----
```


0-NI7	15	3	4.17Kbps	17	3	3.81Kbps	154	120
0-NI6	1	0	1.76Kbps	1	0	1.08Kbps	1080	656
0-NI5	1	0	1.76Kbps	1	0	1.08Kbps	1080	656
0-NI4	1	0	1.76Kbps	1	0	1.08Kbps	1080	656
0-NI3	1	0	1.76Kbps	1	0	1.08Kbps	1080	656
0-NI2	1	0	1.76Kbps	1	0	1.08Kbps	1080	656
0-NI1	1	0	1.76Kbps	1	0	1.08Kbps	1080	656
0-NI0	4108297	821659	10.05Gbps	1	0	1.08Kbps	1509	656
0-HI31	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI30	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI29	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI28	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI27	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI26	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI25	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI24	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI23	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI22	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI21	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI20	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI19	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI18	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI17	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI16	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI14	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI13	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI12	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI11	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI10	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI9	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI8	24556087	4911217	3.30Gbps	4094470	818894	9.95Gbps	64	1500
0-HI6	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412

0-HI5	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI4	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI3	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI2	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI1	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI0	24560241	4912048	3.30Gbps	4095156	819031	9.95Gbps	64	1500

Отбрасывания

Существуют отбрасывания, так как хост не настроен для управления потоками

Шаг 2: С flowcontrol, включенным на хостах, хосты соблюдают паузу и возвращают к исходному состоянию.

```

esc-6001# clear counters ; clear qos stat
esc-6001#
esc-6001# show interface ethernet 147/1/1, ethernet 147/1/9 | inc Ethernet14|rate|pause
Ethernet147/1/1 is up
 30 seconds input rate 4926871976 bits/sec, 410572 packets/sec
 30 seconds output rate 1288637816 bits/sec, 2516870 packets/sec
   input rate 4.93 Gbps, 410.57 Kpps; output rate 1.29 Gbps, 2.52 Mpps
   0 Rx pause
   88129183 Tx pause
Ethernet147/1/9 is up
 30 seconds input rate 4924820632 bits/sec, 410401 packets/sec
 30 seconds output rate 1287225224 bits/sec, 2514111 packets/sec
   input rate 4.92 Gbps, 410.40 Kpps; output rate 1.29 Gbps, 2.51 Mpps
   0 Rx pause
   88069874 Tx pause
esc-6001# show port-channel traffic interface port-channel 147
ChanId      Port Rx-Ucst Tx-Ucst Rx-Mcst Tx-Mcst Rx-Bcst Tx-Bcst
-----
 147  Eth1/41 99.99%  0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/42  0.0%  0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/43  0.0%  0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/44  0.0%  0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/45  0.0%  0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/46  0.0%  0.0% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/47  0.00% 99.00% 12.50% 12.50%  0.0%  0.0%
 147  Eth1/48  0.0%  1.00% 12.50% 12.50%  0.0%  0.0%

```

```

esc-6001# attach fex 147
Attaching to FEX 147 ...
To exit type 'exit', to abort type '$.'
fex-147# dbgexec w

```

```

woo> rate

```

Port	Tx Packets	Tx Rate	Tx Bit	Rx Packets	Rx Rate	Rx Bit	Avg Pkt	Avg
Pkt	(pkts/s)	Rate	(Tx)	(pkts/s)	Rate	(Tx)		
0-NI8	32	6	19.76Kbps	19	3	16.01Kbps	366	506

0-NI7	13	2	3.85Kbps	20	4	5.14Kbps	165	140
0-NI6	1	0	1.76Kbps	2	0	2.16Kbps	1080	656
0-NI5	1	0	1.76Kbps	2	0	2.16Kbps	1080	656
0-NI4	1	0	1.76Kbps	2	0	2.16Kbps	1080	656
0-NI3	1	0	1.76Kbps	2	0	2.16Kbps	1080	656
0-NI2	1	0	1.76Kbps	2	0	2.16Kbps	1080	656
0-NI1	1	0	1.76Kbps	2	0	2.16Kbps	1080	656
0-NI0	4105292	821058	10.04Gbps	2	0	2.16Kbps	1509	656
0-HI31	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI30	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI29	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI28	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI27	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI26	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI25	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI24	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI23	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI22	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI21	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI20	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI19	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI18	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI17	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI16	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI14	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI13	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI12	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI11	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI10	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI9	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI8	12556848	2511369	1.68Gbps	2049754	409950	4.98Gbps	63	1500
0-HI6	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412

0-HI5	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI4	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI3	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI2	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI1	1	0	2.28Kbps	1	0	2.28Kbps	1412	1412
0-HI0	12573036	2514607	1.68Gbps	2051092	410218	4.98Gbps	64	1499

Исправление

По умолчанию, для IP - трафика, FEX распределит нагрузку на основе MAC/IP пары источник/получатель. Для проблем как это, изменяя алгоритм хеширования может помочь получать лучшее распределение трафика по оптоволоконному port-channel. Используйте этот метод, если вы видите неравное распределение нагрузки. Эта опция не является абсолютным решением, поскольку тонкая настройка хэша могла также сделать вещи

```
esc-6001# show port-channel load-balance
```

```
Port Channel Load-Balancing Configuration:
System: source-dest-ip
```

```
Port Channel Load-Balancing Addresses Used Per-Protocol:
Non-IP: source-dest-mac
IP: source-dest-ip source-dest-mac
```

Which hashing algorithm to choose depends on traffic profile. Here are the options available.

```
esc-6001# conf t
```

```
Enter configuration commands, one per line. End with CNTL/Z.
```

```
esc-6001(config)# port-channel load-balance ethernet ?
destination-ip      Destination IP address
destination-mac     Destination MAC address
destination-port    Destination TCP/UDP port
source-dest-ip      Source & Destination IP address (includes 12)
source-dest-ip-only Source & Destination IP addresses only
source-dest-mac     Source & Destination MAC address
source-dest-port    Source & Destination TCP/UDP port (includes 12 and 13)
source-dest-port-only Source & Destination TCP/UDP port only
source-ip           Source IP address
source-mac          Source MAC address
source-port         Source TCP/UDP port
```

Заключения и оптимальные методы

1) Пауза Tx является обычным в рабочем состоянии механизмом для предотвращения отбрасывания пакета в 2232/2248UPQ/B22 FEX

2) Максимизируйте количество каналов связи между 2232/2248UPQ/B22 FEX и родителем. Это должно иметь способность иметь больше путей к сети. Также это поможет в наличии, макс. буферизует для трафика N2H.

- 3) Если каналы связи между FEX и родителем и не равномерно используемым, изменяющимся хешированием port-channel могут помочь.
- 4) С тех пор на FEX нет никакого локального коммутатора, избегайте, имеют профили трафика восток - запад на хостах на FEX.
- 5) Избегайте иметь пульсирующие устройства, такие как устройства NAS, блейд-шасси на FEXes. Они должны быть на родителе.
- 6) Более новый 2348UPQ FEX с 32М совместно используемый буфер, имеет 1 МБ разделенный буфер на NIF для трафика H2N для лучше пакетного поглощения. Также с 40G каналы связи NIF, вероятность перегрузки коллизий/канала связи хэша сведена к минимуму значительно.