

排除在连结2232的TX暂停故障

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[入口缓冲区](#)

[Flow-control配置](#)

[TX暂停的原因在连结2232](#)

[实验室测试个案](#)

[网络图](#)

[测试与在主机没启用的流控制的1.突发数据流](#)

[有在主机的流控制功能的测试2.突发数据流](#)

[测试3.以太网信道哈希冲突](#)

[修正](#)

[结论和最佳实践](#)

简介

本文描述信息为了帮助排除在连结的Transmit(Tx)暂停故障2232个主机接口(HIF)端口。它着重在主机的数据流对网络(H2N)方向(自往网络的服务器进来对北部的数据流，南部)。它不包括方案与对主机(N2H)通信流的网络有关。

本文为连结2232结构增量剂(FEX)主要被创作，但是原理应用B22和2248UPQ FEX。

先决条件

要求

Cisco建议您有这些题目知识

- Cisco连结2000系列配置
- Cisco连结6000系列配置

使用的组件

本文档中的信息基于以下软件和硬件版本：

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

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您使用的是真实网络，请确保您已经了解所有命令的潜在影响。

入口缓冲区

连结2232有32个1/10G面对(伊阵)端口的主机面对(HIF)端口的和8个10G网络。

在您深潜到TX暂停问题里，您需要了解在FEX接口前的可用的缓冲区。缓冲区分配到接口/QOS组可以通过此on命令被检查父母交换机：

```
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>

如看到，在默认服务质量(QoS)，为丢弃组数据流(QOS组0)，FEX HIF有151040个字节为了缓冲H2N数据流，并且XOFF阈值是126720个字节。

Flow-control配置

连结2232过度预定到8:1。默认情况下为了避免在H2N方向的信息包丢弃由于超额预订和缓冲区超出，连结2232安排HIF flow-control传送：

```
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

TX的原因在连结2232暂停

当126720个字节XOFF阈值被击中时，连结2232发送往主机的一次TX暂停在HIF。此的常见原因是：

1. 进入FEX的H2N数据流非常突变性，该造成入口缓冲区是充分的填满并且击中了XOFF阈值。
2. 多数FEX配置使用端口通道聚集多个NIFs。TX暂停也被看到的归结于入口缓冲区，充分获得由于EtherChannel在FEX的哈希冲突。当多个HIF端口尝试对出口在单个伊阵外面由于EtherChannel结果，这发生。

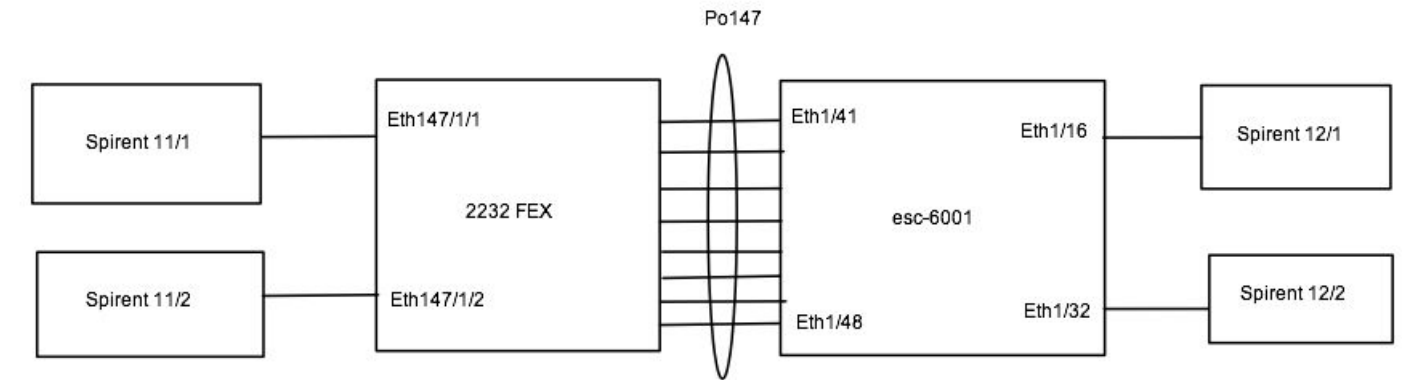
在H2N方向的丢包

TX暂停，一旦XOFF阈值被击中，为了避免信息包丢弃，发送。然而，H2N数据流丢包能被看到是否：

1. 服务器不尊敬暂停或，
2. 有尊敬的延迟造成151KB丢弃阈值被击中的暂停。

实验室测试个案

网络图



对于此测试在实验室里作为主机的，有四个10G spirent端口，两在FEX，并且两在父母连结6001。所有端口在VLAN 50。没有在FEX的其他端口活跃或父母：

```
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
-----
```

测试与在主机没启用的流控制的1.突发数据流

当您发送时100K 1500字节线路费率单播从在Eth147/1/1(to Eth1/16)和Eth147/1/9(do Eth1/32)的主机破裂。每流是单个流。流控制在主机(Spirent)被禁用。

结果：接受器端口报告了大约每流的563被丢弃的数据包。因为流控制在主机被禁用，您能更大量看到TX暂停并且高延迟(大约100微秒钟)：

```
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

```

接受器报告的丢包在FEX实际上下降。有能显示丢包的内部硬件命令，但是要求您完全地了解FEX内部体系结构在本文的范围之外，哪些是。如果需要检查这些计数器，请从事故障排除的此方面的TAC。

有在主机的流控制功能的测试2.突发数据流

当您发送时单个流100K 1500字节线路费率单播从在Eth147/1/1(to Eth1/16)和Eth147/1/9(do Eth1/32)的主机破裂。每流是单个流。流控制在主机(Spirent)被启用。

结果：

接受器端口不报告损失。最小的TX暂停和平均时延是大约19微秒：

```

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
-----
Eth147/1/1   on      on        off      off      0          4739
Eth147/1/9   on      on        off      off      0          4700

```

丢包：

因为主机尊敬从FEX，发送的流控制没有丢包。

测试3.以太网信道哈希冲突

在FEX和父母之间的上行链路是Port-Channel。当它取决于Port-Channel的哪名成员被选时，并且多么繁忙它是，TX暂停在FEX HIFs能被看到。在实验室里，只有两个端口活动在FEX和用于Port-Channel的全部8 uplink端口。

但是对于此测试，与默认散列，数据流从主机在以太网147/1/1和以太网147/1/9获得切细了对连接到Eth1/41在6001的NIF0。如果从主机发送98%线路费率数据流，TX暂停在两被发送HIFs。

对于此测试，流控制在主机被禁用：

```
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 9836009128 bits/sec, 819667 packets/sec
 30 seconds output rate 2516922296 bits/sec, 4915863 packets/sec
   input rate 9.84 Gbps, 819.67 Kpps; output rate 2.52 Gbps, 4.91 Mpps
 0 Rx pause
   98376923 Tx pause
Ethernet147/1/9 is up
 30 seconds input rate 9836252112 bits/sec, 819687 packets/sec
 30 seconds output rate 2516980960 bits/sec, 4915978 packets/sec
   input rate 9.84 Gbps, 819.69 Kpps; output rate 2.52 Gbps, 4.91 Mpps
 0 Rx pause
   98376916 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		(pkts/s)	Rate	(Tx)
(Rx) Err							
0-NI8	24	4	11.23Kbps	22	4	16.49Kbps	272
448							
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									

丢包：

因为主机没有为流控制，被配置有丢包。

有在主机、主机荣誉称号暂停和节流孔返回的flow-control功能：

```
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    ||           | (pkts/s) | Rate    | (Tx) |
(Rx) |Err| |          |          |         ||           |          |         |      |
+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+
| 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
```


修正

默认情况下，对于IP数据流，FEX负载均衡根据源-目的地MAC/IP。对于象这样的问题，请更改Hash算法为了获得数据流的一个更好的分配在结构Port-Channel的。如果看到不同等的负载均衡，请使用此方法。此选项不是一个绝对解决方案：

```
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
```

结论和最佳实践

1. TX暂停是正常可操作的机制为了避免在2232/2248UPQ/B22 FEX的信息包丢弃。
2. 最大化uplink端口的编号在2232/2248UPQ/B22 FEX和父母之间的。为了能有往网络的更多路径并且它帮助有N2H数据流的最大缓冲区。
3. 如果不均匀地使用的uplink端口在FEX和父母之间和，Port-Channel切细的更改可帮助。
4. 因为没有在FEX的本地交换，请避免有在主机的东西方通信流配置文件在FEX。
5. 避免突变性工具例如NAS设备，在FEXes的前端机箱。这些需要在父母。
6. 与32M共有的缓冲区的更新的2348UPQ FEX，有1MB每HIF的共有的缓冲区更好的突发传输吸收的H2N数据流的。并且，与40G伊阵uplink端口，机会的哈希冲突/上行链路拥塞大大减少。