

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[入口缓冲区](#)

[Tx暂停的原因在连结2232](#)

[实验室测试案例](#)

[测试 1：突发数据流用在主机没启用的流量控制。](#)

[测试 2：突发数据流用在主机启用的流量控制](#)

[测试3：以太网信道哈希冲突](#)

[结论和最佳实践](#)

[相关的思科支持社区讨论](#)

简介

本文将提供信息帮助排除故障在连结的TX暂停2232个HIF (主机接口)端口。它着重在H2N (主机对网络)方向(来自往网络，南的服务器的流量的流量到北部)。它不覆盖与N2H (对主机的网络)通信流涉及的案例。

本文为连结主要被创作2232个FEX，但是原理为B22和2248UPQ FEX (结构扩展器)应用。

先决条件

要求

思科建议您有这些主题基础知识

- 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：

- 思科连结N2K-C2232PP-10GE
- 思科连结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 HIF (面对的主机)端口和8个10G伊阵(面对的网络)端口。

在深潜水到Tx暂停问题前，您需要有对在Fex接口的可用的缓冲区的了解。缓冲区分配到接口/Qos类别可以通过以下on命令被检查parent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>

如上所见，与默认服务质量(QoS)，为丢弃类流量(Qos类别0)，FEX HIF有缓冲151040个的字节H2N流量，并且XOFF阈值是126720个字节。

Flow-control配置

连结2232是订购过量8:1。默认情况下要避免在H2N方向的丢包由于超额预订和缓冲超出，连结2232有HIF flow-control发送。

```
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阈值点击，连结2232将发送往主机的一次Tx暂停在HIF。此的常见原因是

1)H2N进入FEX的流量非常突变性造成入口缓冲区是全双工填满和点击XOFF阈值。

2)Most FEX，部署使用端口通道聚集多个NIFs。Tx暂停也被看到的归结于获得在FEX的入口缓冲区全双工由于EtherChannel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

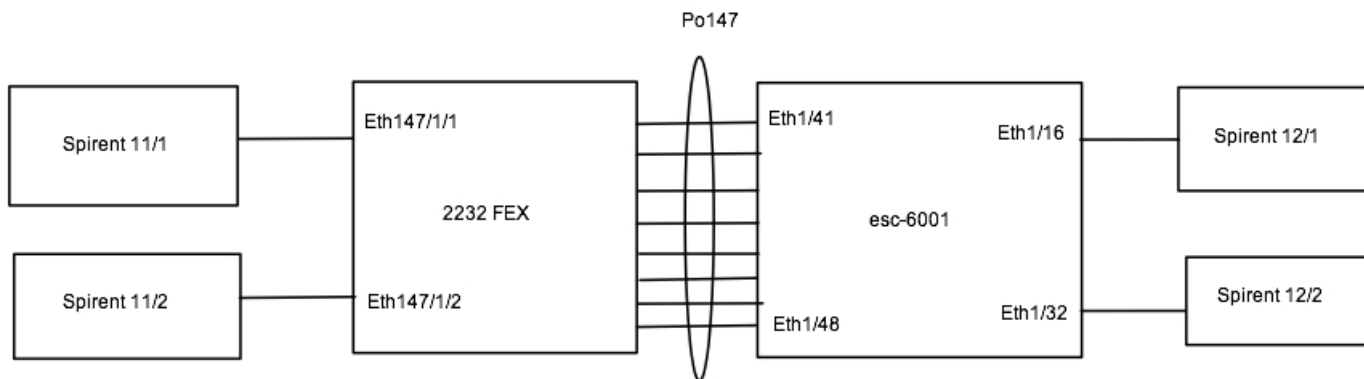
一旦XOFF阈值点击，要避免丢包，Tx暂停发送。然而，H2N流量丢包是否能被看到

1)the服务器不尊敬暂停

2)or有延迟在尊敬造成151KB丢弃阈值的暂停点击。

实验室测试案例

拓扑



对于此测验在实验室里，有作为主机的四个10G spirent端口，两在FEX，并且两在parent连结6001。所有端口在VLAN 50。没有在FEX的其他端口活跃或parent。

```
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 1：突发数据流用在主机没启用的流量控制。

发送100K 1500字节“线路速率”单播从在Eth147/1/1(to Eth1/16)的主机突发传输和Eth147/1/9(do Eth1/32)。每数据流是单个流。流量控制在host(Spirent)禁用。

结果：接收方端口报告大约每个流的563被丢弃的数据包。因为流量控制在主机禁用，您看到批次多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
-----
```

接收方报告的丢包实际上是在fex的dropped。有能显示丢包的内部hardware命令，但是将要求对fex内部体系结构的完整了解在本文的范围之外，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。

测试 2 : 突发数据流用在主机启用的流量控制

发送单个流100K 1500字节“线路速率”单播从在Eth147/1/1(to Eth1/16)的主机突发传输和Eth147/1/9(do Eth1/32)。 每数据流是单个流。 流量控制在host(Spirent)启用。

结果 : 接收方端口不报告损耗。 最小Tx暂停和平均lateny是大约19微妙

```
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，发送的流量控制没有丢包

测试3 : 以太网信道哈希冲突

在FEX和parent之间的上行链路是Port-Channel。根据Port-Channel的哪个成员被选，并且多么忙碌是，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。

但是对于此测验，与默认散列，从主机的流量在以太网147/1/1和以太网147/1/9获得切细了对连接对在6001的Eth1/41的NIF0。如果发送98%线路从主机的速率流量，Tx暂停在两将发送HIFs。

对于此测验流量控制在主机禁用。

```
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 448	24	4	11.23Kbps	22	4	16.49Kbps	272
0-NI7 120	15	3	4.17Kbps	17	3	3.81Kbps	154
0-NI6 656	1	0	1.76Kbps	1	0	1.08Kbps	1080
0-NI5 656	1	0	1.76Kbps	1	0	1.08Kbps	1080
0-NI4 656	1	0	1.76Kbps	1	0	1.08Kbps	1080
0-NI3 656	1	0	1.76Kbps	1	0	1.08Kbps	1080
0-NI2 656	1	0	1.76Kbps	1	0	1.08Kbps	1080
0-NI1 656	1	0	1.76Kbps	1	0	1.08Kbps	1080
0-NI0 656	4108297	821659	10.05Gbps	1	0	1.08Kbps	1509
0-HI31 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI30 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI29 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI28 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI27 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI26 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI25 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI24 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI23 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI22 1412	1	0	2.28Kbps	1	0	2.28Kbps	1412
0-HI21 1412	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

因为主机没有为流量控制，配置有丢包

步骤 2：使用在主机启用的flow-control，主机荣誉称号暂停和节流孔返回。

```

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
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							

修正

默认情况下，对于IP数据流，FEX将负载平衡基于源-目的地MAC/IP。对于象这样的问题，更改哈希算法可帮助获得流量的一更加好的分配在结构Port-Channel的。如果看到不同等的负载均衡，请使用此方法。因为调整哈希可能也做事，此选项不是一绝对解决方案

```
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暂停是避免在2232/2248UPQ/B22 FEX的丢包的正常可操作的机制

2)Maximise uplink端口编号在2232/2248UPQ/B22 FEX和parent之间的。这是有能力有往网络的更多路径。并且它在有将帮助N2H流量的最大缓冲区。

3)If不均匀地使用的uplink端口在FEX和parent之间和，更改Port-Channel切细可帮助。

那里4)Since是在FEX的没有本地交换，避免有在主机的东西方通信流配置文件在FEX。

有的5)Avoid突变性设备例如NAS设备，在FEXes的刀片机箱。这些需要在parent。

6)Newer 2348UPQ FEX用32M共享的缓冲区，有1MB每HIF的共享的缓冲区更加好的突发流量吸收的H2N流量的。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。