

Nexus 2232でのTX一時停止のトラブルシューティング

内容

[概要](#)

[前提条件](#)

[要件](#)

[使用するコンポーネント](#)

[入力バッファ](#)

[フロー制御の設定](#)

[Nexus 2232でのTX一時停止の原因](#)

[ラボテストケース](#)

[ネットワーク図](#)

[テスト1.フロー制御がホストで有効になっていないバーストトラフィック](#)

[テスト2.ホストでフロー制御が有効になっているバーストトラフィック](#)

[テスト3.イーサチャネルハッシュの衝突](#)

[修復](#)

[結論とベストプラクティス](#)

概要

このドキュメントでは、Nexus 2232ホストインターフェイス(HIF)ポートで送信(TX)の一時停止をトラブルシューティングするために役立つ情報について説明します。ホストからネットワーク(H2N)方向のトラフィック(サーバからネットワークに向かう南から北に向かうトラフィック)に焦点を当てます。ネットワークからホスト(N2H)へのトラフィックフローに関連するシナリオについては説明しません。

このドキュメントは主にNexus 2232ファブリックエクステンダ(FEX)用に作成されていますが、この概念はB22および2248UPQ FEXに適用されます。

前提条件

要件

次の項目に関する知識が推奨されます。

- Cisco Nexus 2000 シリーズの設定
- Cisco Nexus 6000 シリーズの設定

使用するコンポーネント

このドキュメントの情報は、次のソフトウェアとハードウェアのバージョンに基づいています。

- Cisco Nexus N2K-C2232PP-10GE

- Cisco Nexus 6001
- 7.1(1)N1(1)

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されました。このドキュメントで使用するすべてのデバイスは、初期（デフォルト）設定の状態から起動しています。対象のネットワークが実稼働中である場合には、どのようなコマンドについても、その潜在的な影響について確実に理解しておく必要があります。

入力バッファ

Nexus 2232には、32個の1/10Gホスト側(HIF)ポートと8個の10Gネットワーク側(NIF)ポートがあります。

TXポーズの問題を詳しく調べる前に、FEXインターフェイスで使用可能なバッファを理解する必要があります。インターフェイス/qos-groupに割り当てられたバッファは、親スイッチで次のコマンドを使用して確認できます。

```
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>

上記のように、デフォルトのQuality of Service(QoS)では、廃棄クラストラフィック(qos-group 0)に対して、H2NトラフィックをバッファするためにFEX HIFに151040バイトがあり、XOFFしきい値は126720バイト0です。

フロー制御の設定

Nexus 2232は8:1にオーバーサブスクライブされています。オーバーサブスクリプションとバッファオーバーランによるH2N方向のパケットドロップを回避するため、Nexus 2232ではデフォルトでHIFフロー制御がオンになっています。

```
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
```

Nexus 2232でのTX一時停止の原因

126720バイトのXOFFしきい値に達すると、Nexus 2232はHIF上のホストに向けてTX Pauseを送信します。一般的な原因は次のとおりです。

1. FEXに着信するH2Nトラフィックは非常にバースト性が高いため、入力バッファがいっぱいになり、XOFFしきい値に達します。
- 2.ほとんどのFEX導入では、複数のNIFを集約するためにポートチャネルを使用します。TX Pauseは、入力バッファが原因で発生し、FEXでのetherchannelハッシュの衝突によりフルになります。これは、Etherchannelの結果が原因で、複数のHIFポートが単一のNIFから出力を試みるときに発生します。

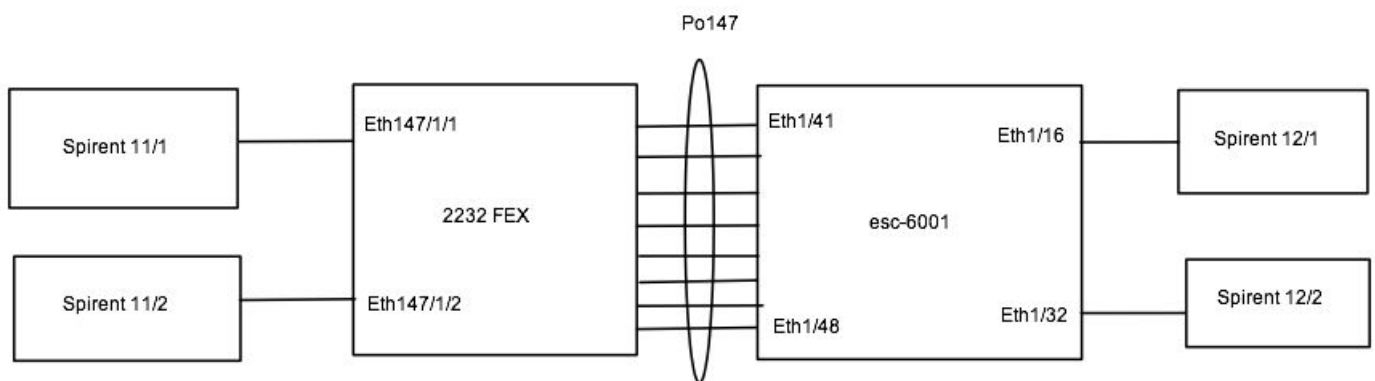
H2N方向でのドロップ

パケットドロップを回避するために、XOFFしきい値に達するとTX Pauseが送信されます。ただし、H2Nトラフィックのドロップは、次の場合に発生します。

1. サーバは一時停止または
- 2.151KBのドロップしきい値に達する原因となる一時停止を維持するために遅延を設けます。

ラボテストケース

ネットワーク図



このラボのテストでは、ホストとして機能する4つの10G spirentポートがあり、2つはFEXにあり、2つは親Nexus 6001にあります。すべてのポートがVLAN 50にあります。FEXまたは親でアクティブなポートは他にありません。

```
esc-6001# show port-channel summary
Flags:  D - Down          P - Up in port-channel (members)
        I - Individual    H - Hot-standby (LACP only)
        s - Suspended     r - Module-removed
        S - Switched     R - Routed
        U - Up (port-channel)
        M - Not in use. Min-links not met
```

Group	Port-Channel	Type	Protocol	Member Ports
147	Po147(SU)	Eth	NONE	Eth1/41(P) Eth1/42(P) Eth1/43(P) Eth1/44(P) Eth1/45(P) Eth1/46(P) Eth1/47(P) Eth1/48(P)

```

esc-6001# show fex 147 detail | exc Down
FEX: 147 Description: FEX0147    state: Online
FEX version: 7.1(1)N1(1) [Switch version: 7.1(1)N1(1)]
FEX Interim version: 7.1(1)N1(1)
Switch Interim version: 7.1(1)N1(1)
Extender Serial: FOT1635R003
Extender Model: N2K-C2232PP-10GE, Part No: 73-12533-05
Card Id: 82, Mac Addr: 20:3a:07:34:5b:02, Num Macs: 64
Module Sw Gen: 21 [Switch Sw Gen: 21]
post level: complete
Pinning-mode: static    Max-links: 1
Fabric port for control traffic: Eth1/47
FCoE Admin: false
FCoE Oper: true
FCoE FEX AA Configured: false
Fabric interface state:
  Po147 - Interface Up. State: Active
  Eth1/41 - Interface Up. State: Active
  Eth1/42 - Interface Up. State: Active
  Eth1/43 - Interface Up. State: Active
  Eth1/44 - Interface Up. State: Active
  Eth1/45 - Interface Up. State: Active
  Eth1/46 - Interface Up. State: Active
  Eth1/47 - Interface Up. State: Active
  Eth1/48 - Interface Up. State: Active
Fex Port      State Fabric Port
  Eth147/1/1   Up      Po147
  Eth147/1/9   Up      Po147

```

Logs:

```

04/21/2015 21:58:30.162193: Module register received
04/21/2015 21:58:30.164611: Registration response sent
04/21/2015 21:58:30.196708: create module inserted event.
04/21/2015 21:58:30.197425: Module Online Sequence
04/21/2015 21:58:35.051474: Module Online

```

テスト1.フロー制御がホストで有効になっていないバーストトラフィック

Eth147/1/1(Eth1/16)およびEth147/1/9(do Eth1/32)のホストから100K 1500バイトのラインレートユニキャストバーストを送信する場合。各ストリームは1つのフローです。ホスト(Spirent)でフロー制御が無効になっている。

[Results] : 受信側のポートで、各フローに対して約563個のパケットが廃棄されたことが報告されました。フロー制御がホストで無効になっているため、TXの一時停止が大幅に増加し、遅延も大きくなります (約100マイクロ秒)。

```

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

レシーバによって報告されたドロップは、実際にはFEX自体でドロップされます。ドロップを表示できる内部ハードウェアコマンドがありますが、このドキュメントの範囲外であるFEX内部アーキテクチャを完全に理解する必要があります。これらのカウンタを確認する必要がある場合は、トラブルシューティングのこの側面についてTACに連絡してください。

テスト2.ホストでフロー制御が有効になっているバーストトラフィック

Eth147/1/1(Eth1/16)およびEth147/1/9(do Eth1/32)のホストから単一フロー100K 1500バイトのラインレートユニキャストバーストを送信する場合。各ストリームは1つのフローです。ホスト(Spirent)でフロー制御が有効になっている。

[Results] :

受信側のポートでは損失が報告されません。最小のTX一時停止と平均遅延は約19マイクロ秒です。

```
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

廃棄 :

ホストはFEXから送信されたフロー制御を優先するため、ドロップはありません。

テスト3.イーサチャネルハッシュの衝突

FEXと親の間のアップリンクはポートチャネルです。ポートチャネル内のどのメンバが選択され、どのメンバがビジーであるかによって異なりますが、FEX HIFでTX Pauseを確認できます。ラボでは、FEXでアクティブなポートが2つだけで、ポートチャネルで使用される8つのアップリンクはすべてアクティブです。

ただし、このテストでは、デフォルトのハッシュを使用して、イーサネット147/1/1およびイーサネット147/1/9のホストからのトラフィックは、6001のEth1/41に接続するNIF0にハッシュされま

す。ホストから98%のラインレートトラフィックを送信すると、両方のHIFでTX Pauseが送信されます。

このテストでは、ホストでフロー制御が無効になっています。

```
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

ホストがフロー制御用に設定されていないため、ドロップが発生します。

ホストでフロー制御が有効になっている場合、ホストは一時停止とスロットルバックを実行し
ます。

```
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									

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

修復

IPトラフィックの場合、デフォルトでは、FEXロードバランシングは発信元と宛先のMAC/IPに基づいています。このような問題の場合は、ファブリックポートチャンネル上でトラフィックをより適切に分散するためにハッシュアルゴリズムを変更します。不等ロードバランシングが発生する場合は、この方法を使用します。このオプションは絶対的なソリューションではありません。

```
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
```

結論とベストプラクティス

1. TX Pauseは、2232/2248UPQ/B22 FEXでのパケットドロップを回避するための正常な動作メカニズムです。
2. 2232/2248UPQ/B22 FEXと親の間のアップリンク数を最大にします。ネットワークへのパスを増やすことができ、N2Hトラフィックの最大バッファを確保することもできます。
3. FEXと親の間のアップリンクが均等に使用されていない場合は、ポートチャンネルハッシングの変更が役立ちます。
4. FEXにはローカルスイッチングがないため、FEX上のホストに水平方向のトラフィックフロープロファイルを設定しないでください。
5. FEX上のNASデバイス、ブレードシャーシなどのバーストアプライアンスを避けます。これらは親である必要があります。
6. 新しい2348UPQ FEXと32Mの共有バッファ、H2Nトラフィック用のHIFあたり1MB共有バッファを備え、バースト吸収を改善します。また、40G NIFアップリンクでは、ハッシュの衝突やア

アップリンクの輻輳の可能性が大幅に低減されます。