解決由分散式Etherchannel引起的介面溢位

目錄

簡介

EtherChannel捆綁用於提供高頻寬互連。本文討論適用於執行Supervisor 720（搭載PFC3A、PFC3B或PFC3BXL）的Catalyst 6500交換器上的Cisco EtherChannel的限制，這可能會導致Etherchannel成員介面上的溢位增加。此限制與第2層轉發引擎相關，因此僅適用於第2層EtherChannel。

必要條件

需求

本文件沒有特定需求。

採用元件

本文中的資訊是根據執行Supervisor Engine 720的Cisco Catalyst 6500系列交換器。本實驗設定已使用WS-X6704-10GE。WS-X6704-10GE是一個Catalyst 6500模組，沒有超額訂閱，且具有2個交換矩陣通道連線，每個連線20 Gbps。

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路正在作用，請確保您已瞭解任何指令可能造成的影響。

問題

設定第2層分散式EtherChannel(DEC)時，Catalyst 6500可能會遇到介面溢位。DEC是etherchannel 2個或更多分散式轉發卡(DFC)模組。具有兩個成員介面的EtherChannel連線到同一線卡但連線在不同的交換矩陣通道上，不視為DEC。

溢位計數器帳戶接收方硬體無法將接收的資料交給硬體緩衝區的次數。換句話說，流量的輸入速率超過了接收器處理資料的能力。

在某些情況下，DFC提供重新循環資料包的功能。循環可用於在ACL或QoS三重內容可定址儲存器(TCAM)、NetFlow表或轉發資訊庫(FIB)TCAM表中執行額外的查詢。 封包重新循環僅發生在特定封

包流程上；其他資料包流不受影響。在模組上發生資料包的重寫；然後將封包轉回原則功能卡(PFC)以進行其他處理。

當使用第2層DEC時，需要在 資料包轉發。 在以下情況下，多模組L2 EtherChannel也需要重新循環 Catalyst 6500與3B/3BXL PFC模式一起以直通模式運行。

有關直通模式的詳細資訊，請參閱。

當交換矩陣利用率達到約50%時，溢位計數器可能會開始遞增。

疑難排解與驗證

1)查詢Etherchannel中遇到遞增超限的成員介面。

```
6500#show etherchannel summary
Flags:  D - down        P - bundled in port-channel
        I - stand-alone s - suspended
        H - Hot-standby (LACP only)
        R - Layer3      S - Layer2
        U - in use      N - not in use, no aggregation
        f - failed to allocate aggregator
        M - not in use, no aggregation due to minimum links not met
        m - not in use, port not aggregated due to minimum links not met
        u - unsuitable for bundling
        d - default port
        w - waiting to be aggregated

Number of channel-groups in use: 2
Number of aggregators:           2
Group  Port-channel  Protocol    Ports
------+-------------+-----------+-----------------------------------------------
10     Po10(SU)         -        Te2/1(P)       Te3/1(P)
20     Po20(SU)         -        Te2/2(P)       Te3/2(P)
```

2)驗證成員介面上的輸入速率和溢位計數器。

```
6500#show interfaces tenGigabitEthernet 2/1
TenGigabitEthernet2/1 is up, line protocol is up (connected)
Hardware is C6k 10000Mb 802.3, address is 0002.fcc1.21ac (bia 0002.fcc1.21ac)
MTU 1500 bytes, BW 10000000 Kbit, DLY 10 usec,
reliability 255/255, txload 1/255, rxload 251/255
Encapsulation ARPA, loopback not set
Keepalive set (10 sec)
Full-duplex, 10Gb/s, media type is 10Gbase-SR
input flow-control is on, output flow-control is off
ARP type: ARPA, ARP Timeout 04:00:00
Last input never, output 00:00:51, output hang never
Last clearing of "show interface" counters 00:00:13
Input queue: 0/2000/5597178/0 (size/max/drops/flushes); Total output drops: 0
Queueing strategy: fifo
Output queue: 0/40 (size/max)
30 second input rate 9868906000 bits/sec, 822409 packets/sec
30 second output rate 3000 bits/sec, 5 packets/sec
10968368 packets input, 16452552000 bytes, 0 no buffer
Received 0 broadcasts (0 multicasts)
0 runts, 0 giants, 0 throttles
0 input errors, 0 CRC, 0 frame, 5597195 overrun, 0 ignored
```

```
0 watchdog, 0 multicast, 0 pause input
0 input packets with dribble condition detected
79 packets output, 5596 bytes, 0 underruns
0 output errors, 0 collisions, 0 interface resets
0 babbles, 0 late collision, 0 deferred
0 lost carrier, 0 no carrier, 0 PAUSE output
0 output buffer failures, 0 output buffers swapped out

6500#show interfaces tenGigabitEthernet 2/2
TenGigabitEthernet2/2 is up, line protocol is up (connected)
Hardware is C6k 10000Mb 802.3, address is 0002.fcc1.21ad (bia 0002.fcc1.21ad)
MTU 1500 bytes, BW 10000000 Kbit, DLY 10 usec,
reliability 255/255, txload 1/255, rxload 251/255
Encapsulation ARPA, loopback not set
Keepalive set (10 sec)
Full-duplex, 10Gb/s, media type is 10Gbase-SR
input flow-control is on, output flow-control is off
ARP type: ARPA, ARP Timeout 04:00:00
Last input never, output 00:00:26, output hang never
Last clearing of "show interface" counters 00:00:03
Input queue: 0/2000/45043/0 (size/max/drops/flushes); Total output drops: 0
Queueing strategy: fifo
Output queue: 0/40 (size/max)
30 second input rate 9868149000 bits/sec, 822345 packets/sec
30 second output rate 2000 bits/sec, 4 packets/sec
8823464 packets input, 13233496000 bytes, 0 no buffer
Received 0 broadcasts (0 multicasts)
0 runts, 0 giants, 0 throttles
0 input errors, 0 CRC, 0 frame, 4575029 overrun, 0 ignored
0 watchdog, 0 multicast, 0 pause input
0 input packets with dribble condition detected
0 packets output, 0 bytes, 0 underruns
0 output errors, 0 collisions, 0 interface resets
0 babbles, 0 late collision, 0 deferred
0 lost carrier, 0 no carrier, 0 PAUSE output
0 output buffer failures, 0 output buffers swapped out
```

3)找出存在這些介面的模組。

```
6500#show module
Mod Ports Card Type                                  Model              Serial No.
--- ----- -------------------------------------- ------------------ -----------
 2    4  CEF720 4 port 10-Gigabit Ethernet      WS-X6704-10GE      SAD07430301
 3    4  CEF720 4 port 10-Gigabit Ethernet      WS-X6704-10GE      SAL1316NJD4
 5    2  Supervisor Engine 720 (Active)         WS-SUP720-3B       JAF1224BFSQ
Mod MAC addresses                       Hw    Fw           Sw           Status
--- ---------------------------------- ------ ------------ ------------ -------
 2  0002.fcc1.21ac to 0002.fcc1.21af   1.2   12.2(14r)S5  12.2(33)SXI5 Ok
 3  0024.c4f5.b2f4 to 0024.c4f5.b2f7   2.9   12.2(14r)S5  12.2(33)SXI5 Ok
 5  001e.be6e.99b4 to 001e.be6e.99b7   5.6   8.5(4)       12.2(33)SXI5 Ok
Mod  Sub-Module                  Model              Serial       Hw     Status
---- --------------------------- ------------------ ----------- ------- -------
 2  Centralized Forwarding Card WS-F6700-CFC       SAD0743039A  1.1    Ok
 3  Distributed Forwarding Card WS-F6700-DFC3B     SAL1408BP0Y  4.8    Ok
 5  Policy Feature Card 3       WS-F6K-PFC3B       JAF1223BAPB  2.3    Ok
 5  MSFC3 Daughterboard         WS-SUP720          JAF1223BACM  3.1    Ok
Mod  Online Diag Status
---- -------------------
 2  Pass
 3  Pass
 5  Pass
```

4)找出與這些模組對應的交換矩陣介面利用率。

```
6500#show fabric utilization
slot    channel      speed    Ingress %     Egress %
   2          0        20G            0            0
   2          1        20G           49            0
   3          0        20G            0            0
   3          1        20G            0           50
   5          0        20G            0            0
```

5)如上所述，介面Tengigabitethernet 2/1和Tengigabitethernet 2/2以線速運行，但由於入口流量控制，入口處會發生溢位（丟棄）。為了檢查交換機是否正在經歷流量控制並且正運行到上述限制，請使用以下命令。

```
6500#show platform hardware capacity rewrite-engine performance slot 2
slot channel perf_id description          packets              total overruns
----+-------+-------+--------------------+--------------------+--------------+
2    0       0x235   FAB RX 0             41083                0
2    0       0x237   FAB RX 1             0                    0
2    0       0x27B   FAB TX 0             1904                 0
2    0       0x27F   FAB TX 1             0                    0
2    0       0x350   REPLICATION ML3      0                    0
2    0       0x351   REPLICATION ML2      0                    0
2    0       0x352   RECIRC L2            0                    0
2    0       0x353   RECIRC L3            0                    0
2    0       0x34C   SPAN TX 0            0                    0
2    0       0x34D   SPAN TX 1            0                    0
2    0       0x34E   SPAN RX 0            0                    0
2    0       0x34F   SPAN RX 1            0                    0
2    0       0x354   SPAN TERMINATION     0                    0
2    1       0x235   FAB RX 0             759500888            0
2    1       0x237   FAB RX 1             0                    0
2    1       0x27B   FAB TX 0             4545890286           0
2    1       0x27F   FAB TX 1             0                    0
2    1       0x350   REPLICATION ML3      0                    0
2    1       0x351   REPLICATION ML2      0                    0
2    1       0x352   RECIRC L2            68615145             1047 <<< L2 Recirculation
2    1       0x353   RECIRC L3            0                    0
2    1       0x34C   SPAN TX 0            0                    0
2    1       0x34D   SPAN TX 1            0                    0
2    1       0x34E   SPAN RX 0            0                    0
2    1       0x34F   SPAN RX 1            0                    0
2    1       0x354   SPAN TERMINATION     0                    0
```

上面的輸出表明入口複製引擎正在執行L2再循環，因為L2 DEC存在。由於存在重新循環，資料包到達複製引擎的時間是原來的兩倍，而頻寬消耗是原來的兩倍。換句話說，單個交換矩陣通道的效能減半，因為每個 每個資料包 在內部見過兩次。

解決方案

1)配置不受此限制影響的非分散式EtherChannel。

為了驗證此理論，在同一模組（非DEC）上的介面上配置了一個EtherChannel，並且觀察到，以與上面相同的資料包速率，介面不會看到任何溢位遞增。這可避免此問題。

```
6500#show fabric utilization
slot    channel      speed    Ingress %     Egress %
   2          0        20G            0            0
   2          1        20G           99            0
   3          0        20G            0            0
```

```
   3          1        20G            0           99
   5          0        20G            0            0

TenGigabitEthernet2/1 is up, line protocol is up (connected)
 Hardware is C6k 10000Mb 802.3, address is 0002.fcc1.21ac (bia 0002.fcc1.21ac)
 MTU 1500 bytes, BW 10000000 Kbit, DLY 10 usec,
    reliability 255/255, txload 1/255, rxload 251/255
 Encapsulation ARPA, loopback not set
 Keepalive set (10 sec)
 Full-duplex, 10Gb/s, media type is 10Gbase-SR
 input flow-control is on, output flow-control is off
 ARP type: ARPA, ARP Timeout 04:00:00
 Last input never, output 00:00:06, output hang never
 Last clearing of "show interface" counters 00:36:12
 Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops: 0
 Queueing strategy: fifo
 Output queue: 0/40 (size/max)
 30 second input rate 9868487000 bits/sec, 822374 packets/sec
 30 second output rate 3000 bits/sec, 6 packets/sec
    1783710310 packets input, 2675565466500 bytes, 0 no buffer
    Received 0 broadcasts (0 multicasts)
    0 runts, 0 giants, 0 throttles
    0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
    0 watchdog, 0 multicast, 0 pause input
    0 input packets with dribble condition detected
    13115 packets output, 946206 bytes, 0 underruns
    0 output errors, 0 collisions, 0 interface resets
    0 babbles, 0 late collision, 0 deferred
    0 lost carrier, 0 no carrier, 0 PAUSE output
    0 output buffer failures, 0 output buffers swapped out

TenGigabitEthernet2/2 is up, line protocol is up (connected)
 Hardware is C6k 10000Mb 802.3, address is 0002.fcc1.21ad (bia 0002.fcc1.21ad)
 MTU 1500 bytes, BW 10000000 Kbit, DLY 10 usec,
    reliability 255/255, txload 1/255, rxload 251/255
 Encapsulation ARPA, loopback not set
 Keepalive set (10 sec)
 Full-duplex, 10Gb/s, media type is 10Gbase-SR
 input flow-control is on, output flow-control is off
 ARP type: ARPA, ARP Timeout 04:00:00
 Last input never, output 00:00:11, output hang never
 Last clearing of "show interface" counters 00:37:31
 Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops: 0
 Queueing strategy: fifo
 Output queue: 0/40 (size/max)
 30 second input rate 9868462000 bits/sec, 822371 packets/sec
 30 second output rate 3000 bits/sec, 6 packets/sec
    1849499775 packets input, 2774249662500 bytes, 0 no buffer
    Received 0 broadcasts (0 multicasts)
    0 runts, 0 giants, 0 throttles
    0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
    0 watchdog, 0 multicast, 0 pause input
    0 input packets with dribble condition detected
    13599 packets output, 980928 bytes, 0 underruns
    0 output errors, 0 collisions, 0 interface resets
    0 babbles, 0 late collision, 0 deferred
    0 lost carrier, 0 no carrier, 0 PAUSE output
    0 output buffer failures, 0 output buffers swapped out
```

2)在**需要L2 DEC的情況下，在PFC 3C/3CXL模式下使用Catalyst 6500交換機。**

附註：如果現有模組正在運行DFC3A/ DFC3B/ DFC3BXL，則需要升級DFC硬體。

3)**如果設計和配置適用於CSCti23324中的條件，請升級IOS版本。**

此錯誤修正程式放寬了僅使用具有67xx模組的Catalyst 6500交換器的L2 DEC或多模組EtherChannel的再循環要求。Cisco IOS版本12.2(33)SXJ1和更新版本中已解決此錯誤。 請注意以下適用於此錯誤的要點。

a)Bug fix放寬了L2 DEC或多模組EC的再循環要求 僅適用於Catalyst 6500交換機67xx模組。如果Catalyst 6500交換機至少 跨任何較舊的DFC模組（例如6516/6816）或組合的67xx的一個L2 DEC 和6516/6818模組，所有配置的第2層DEC將強制重新循環 在系統中。如果Catalyst 6500交換機具有任何較舊的模組且已配置 僅在67xx模組上使用L2 DEC，不會強制實施再循環。

b)所有67xx線卡的存在不足以消除DEC的再循環要求。例如，如果您在2 6704 DFC上配置了DEC，在6748 CFC上配置了另一個埠通道，系統將檢查管理引擎的轉發引擎（對於CFC模組），並開始使用再循環。

c)對於VS-SUP720-10G，此錯誤修復程式在CFC線路卡/管理引擎上至少有一個L2 DEC埠的情況下無法正常工作。在這種情況下，仍會發生再循環。此外，即使從埠通道中刪除啟用了Supervisor/CFC的埠，鄰接關係也不會升級，並且重新循環仍然存在。在這種情況下，需要重新載入才能對硬體重新程式設計，而刪除並重新配置埠通道/冗餘切換/刪除第2層VLAN等操作則無濟於事。