

Limitação de distância F1 Multihop de FCoE do nexo 7000

Índice

[Introdução](#)

[Problema](#)

[Solução](#)

[Hardware/software da elevação a fim apoiar F2/F2e](#)

Introdução

Este documento descreve o que fazer se você experimenta descartes da entrada no Fibre Channel sobre relações Multihop dos Ethernet (FCoE). Estes problema/documento da solução é úteis quando os sintomas do descarte são identificados nas relações que interconectam datacenters remotos.

Este exemplo descreve um cenário de vida real desta edição.

A topologia mostrada no exemplo descreve dois datacenters separados por 10KM. Há uma expansão virtual de 10KM FCoE (VE) (multihop) a relação que conecta o DC1 e o DC2. As relações multihop são configuradas nas placas de linha N7K-F132XP-15. Por esta [folha de dados do F1 Series](#), isto deve ter estado dentro do alcance suportado.

Inicialmente, a folha de dados indicou este o centro de dados da IEEE que constrói uma ponte sobre as características (DCB):

- controle de fluxo Prioridade-baseado (PFC): IEEE P802.1Qbb
- Seleção aumentada da transmissão (ET): IEEE P802.1Qaz
- Centro de dados que constrói uma ponte sobre a troca (DCBX)
- Distância de link sem perda máxima: 20 quilômetros

A identificação de bug Cisco [CSCts72420](#) foi alterada a fim endereçar a documentação. A linha com respeito à distância de link sem perda de 20KM foi removida.

Problema

Os dispositivos EMC VPLEX apoiam uma característica da replicação do armazenamento. Esta replicação síncrono usada encenação. Quando os dispositivos EMC VPLEX foram promovidos tornaram-se “fora da sincronização”. A elevação do cargo VPLEX, os dispositivos começou a replicate quantidades elevadas de dados sobre o link multihop de 10KM FCoE.

Quando a replicação dos dados aumentou, estes eventos transpiraram:

1. O nexa 5000-DC2 começou a enviar a pausa constante ao nexa 7000-DC2.
2. O nexa 7000-DC2 começou a enviar a pausa sobre o link de 10KM ao nexa 7000-DC1.
3. O nexa 7000-DC1 enviado pausa ao nexa 5000-DC1, e assim por diante.

Estes eventos são uma visualização de alto nível do comportamento de controle de fluxo previsto de FCoE. Os frames de pausa recebidos do nexa 5000-DC2 indicam congestões em um dispositivo final. Enquanto os buffers de ingresso começam a se encher, os frames de pausa gotejam de novo na tela.

A edição nesta encenação é que o nexa 7000-DC2 rejeitou constantemente pacotes no ingresso sobre o link do multithop de 10KM.

```

Ethernet4/1 is up
Dedicated Interface
Hardware: 1000/10000 Ethernet, address: XXXX.XXXX.XXXX (bia XXXX.XXXX.XXXX)
MTU bytes (CoS values): 9216(0-2,4-7) 2112(3)
BW 10000000 Kbit, DLY 10 usec, reliability 255/255, txload 1/255, rxload 1/255
Encapsulation ARPA
Port mode is trunk
full-duplex, 10 Gb/s, media type is 10G
Beacon is turned off
Auto-Negotiation is turned on
Input flow-control is off, output flow-control is off
Rate mode is dedicated
Switchport monitor is off
EtherType is 0x8100
Last link flapped 25week(s) 0day(s)
Last clearing of "show interface" counters 79w2d
30 seconds input rate 296186536 bits/sec, 27891 packets/sec
30 seconds output rate 151677360 bits/sec, 19294 packets/sec
Load-Interval #2: 5 minute (300 seconds)
input rate 289.58 Mbps, 27.61 Kpps; output rate 165.20 Mbps, 20.05 Kpps
RX
566235497816 unicast packets 2504479 multicast packets 0 broadcast packets
566239834433 input packets 502487779153524 bytes
219280594774 jumbo packets 0 storm suppression packets
0 runts 0 giants 0 CRC 0 no buffer
0 input error 0 short frame 0 overrun 0 underrun 0 ignored
0 watchdog 0 bad etype drop 0 bad proto drop 0 if down drop
0 input with dribble 19312516 input discard
1832141 Rx pause
TX
681040135255 unicast packets 2504251 multicast packets 0 broadcast packets
681046392756 output packets 744942450903588 bytes
333793360248 jumbo packets
0 output error 0 collision 0 deferred 0 late collision
0 lost carrier 0 no carrier 0 babble 0 output discard
3753250 Tx pause
5 interface resets

```

Isto não deve acontecer enquanto a relação acima leva somente o tráfego de FCoE (CoS 3). Os descartes da entrada violam a política de QoS da “nenhum-gota” para FCoE. Além disso, os descartes em um ambiente de FCoE podiam conduzir aos abortos SCSI, erros, e assim por diante.

Solução

Quando um dispositivo envia a pausa, a relação que gerencie o frame de pausa deve ter uma fila de ingresso com um espaço de buffer grande bastante para proteger duas vezes a distância de

link. Isto é porque naquele tempo aquele que a pausa é gerada o fio pôde estar completo. No tempo onde o dispositivo adjacente recebe/processa o frame de pausa gerado, o fio pôde estar completo outra vez. Assim o dispositivo que gerencie a pausa deve ter a capacidade para proteger duas vezes a distância de link.

Em cima do cálculo, poderia ter havido os pacotes 100+ em voo sobre o link de 10KM. Devido a uma limitação de ASIC, a placa de linha do F1 Series não pode apoiar FCoE sem perda em um link de 10KM ou maior.

Nota: O buffer de ingresso (IB) é usado para enfileirar em voo pacotes antes da pausa. Depois que a pausa é enviada, IB está usado já não. O buffer da latência (LB) está usado para enfileirar em voo pacotes depois que a pausa é enviada.

Hardware/software da elevação a fim apoiar F2/F2e

A identificação de bug Cisco [CSCua10484](#) endereçou o apoio sem perda da distância do longo-curso F2. Em NX-OS libere 6.1(2) e mais atrasado, estas alterações de configuração são permitidas.

Nota: Nas versões anterior do código, a mudança é permitida mas não toma o efeito.

O espaço deixado no IB para travar pacotes pode ser calculado como: PL_STOP - PL_PAUSE. À revelia os valores PL_STOP e de HWM (PL_PAUSE) são os mesmos.

```
module-4# show hardware internal mac port 1 qos configuration | begin IB | end EB
IB
Port page limit : 3584 (1376256 Bytes)
VL# HWM pages(bytes) LWM pages(bytes) Used PL_STOP(HWM & LWM)
pages THR
0 1107 ( 425088) 1059 ( 406656) 0 1107 1059
1 2 ( 768) 1 ( 384) 0 2 1
2 1107 ( 425088) 1059 ( 406656) 0 1107 1059
3 1053 ( 404352) 1029 ( 395136) 0 1053 1029
4 2 ( 768) 1 ( 384) 0 2 1
5 231 ( 88704) 159 ( 61056) 0 231 159
6 2 ( 768) 1 ( 384) 0 2 1
7 2 ( 768) 1 ( 384) 0 2 1
Credited DWRR WT: 216 (0xd8) Uncredited DWRR WT: 144 (0x90)
DWRR honor UC = FALSE
Leak Lo weight = 0xd8, enabled = FALSE
EB
```

Você pode alterar estes valores a fim apoiar uma distância maior pela atribuição de maiores buffers ao Classe de serviço (CoS) da nenhum-gota. A fim terminar isto, duplique o mapa de política do Qualidade de Serviço (QoS) 'default-4q-7e-in-policy.

Nota: O cenário de configuração seguinte é costurado para um aumento na atribuição do buffer de ingresso COS 3 no nex 7000 com padrão VDC, não Admin VDC. Desde que o Admin VDC é restritamente um "Gerenciamento" VDC somente, sem a funcionalidade do DATA-plano, as etapas adicionais para copiar a política de QoS e para alterar as alocações de buffer não são precisados, dentro deste VDC. Consequentemente, se você usa o Admin VDC, simplesmente faça estas alterações no armazenamento VDC.

DWRR honor UC = FALSE

Leak Lo weight = 0xd8, enabled = FALSE

EB

No exemplo, o espaço deixado em IB = 3521 páginas - 1161 páginas = => de 2360 páginas
906,240 bytes.

Nota: Esta ação alternativa é aceitável quando SOMENTE o tráfego de FCoE é passado sobre a relação do multihop de FCoE. Se o tráfego de dados é passado também, contacte o centro de assistência técnica da Cisco (TAC) para o auxílio.

Ou

Se disponível, use o Fibre Channel nativo (FC) entre locais. Esta solução exige a intervenção grosseira ou a fibra oculta do Multiplexer da divisão de comprimento de onda/Multiplexer do dense wavelength division (CWDM/DWDM), dependente da distância exigida.