

FlexPod 공통 성능 문제

목차

[소개](#)

[FlexPod 개념 개요](#)

[성능 고려 사항](#)

[환경](#)

[측정](#)

[초기 계획](#)

[FlexPod의 성능 문제](#)

[일반적인 문제](#)

[프레임 및 패킷 손실](#)

[MTU 불일치](#)

[Nexus 5000 및 UCS 플랫폼의 MTU 디스플레이](#)

[엔드 투 엔드 구성](#)

[엔드 투 엔드 점보 프레임 테스트](#)

[버퍼 관련 문제](#)

[드라이버 문제](#)

[어댑터 정보](#)

[논리적 패킷 흐름](#)

[입력/출력 모듈](#)

[설계 고려 사항](#)

[포트 속도 선택 및 포트 채널 고려 사항](#)

[스토리지 관련 문제](#)

[스토리지 배치](#)

[최적 경로 선택](#)

[VM 및 하이퍼바이저 트래픽 공유](#)

[문제 해결 팁](#)

[문제 축소](#)

[Cisco](#)

[카운터 제한 사항](#)

[컨트롤 플레인 고려 사항](#)

[트래픽 캡처](#)

[NetApp](#)

[VMware](#)

[알려진 문제 및 개선 사항](#)

[TAC 케이스](#)

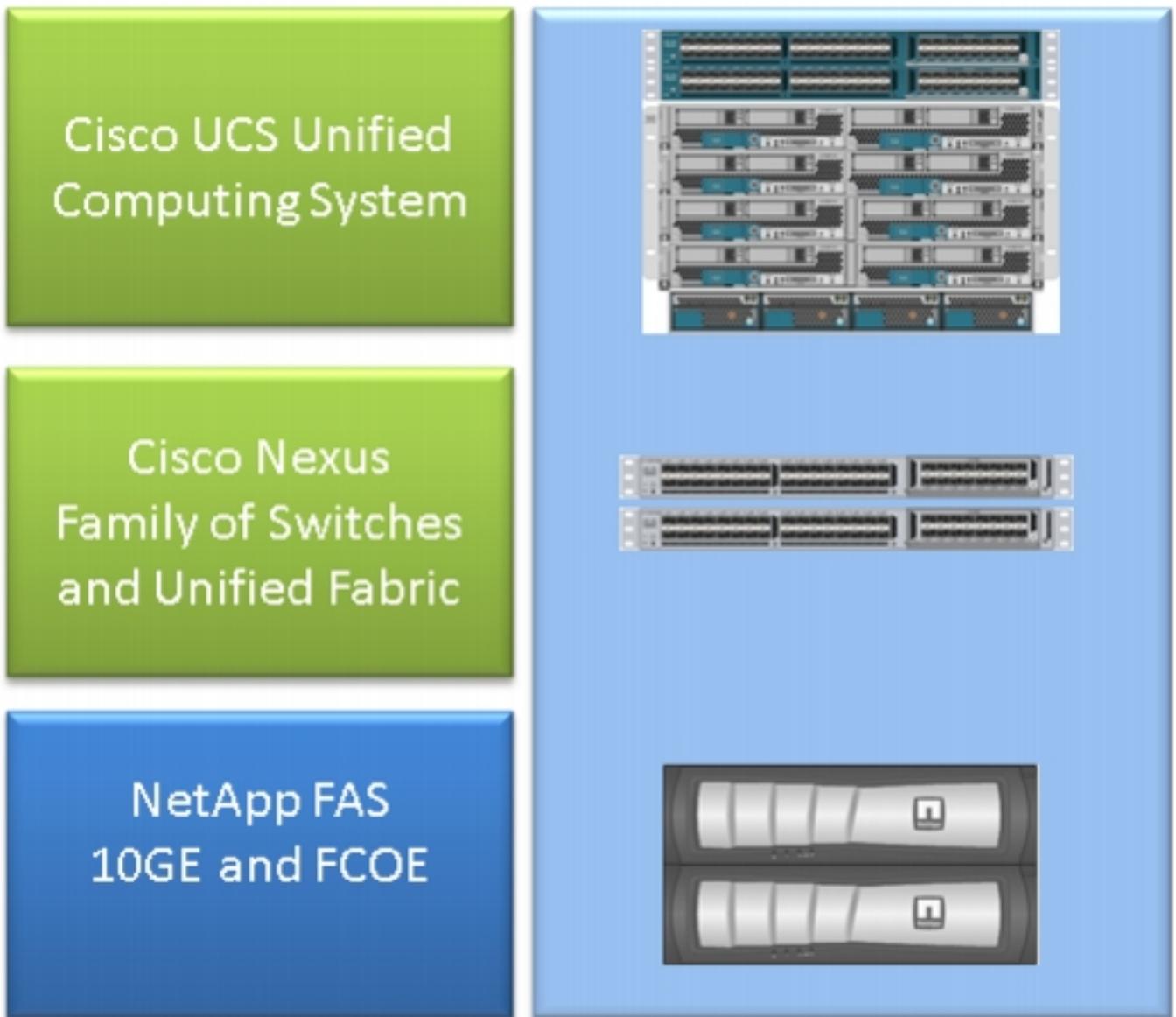
[피드백](#)

소개

이 문서에서는 FlexPod 환경의 일반적인 성능 문제를 설명하고, 문제를 해결하는 방법을 제공하며, 완화 단계를 제공합니다. FlexPod 환경에서 성능 문제를 해결하려는 고객을 위한 시작점으로 사용됩니다. 이 문서는 최근 몇 달 동안 TAC(Data Center Solutions Technical Assistance Center) 팀에서 확인한 문제의 결과로 작성되었습니다.

FlexPod 개념 개요

FlexPod는 Nexus 스위치를 통해 NetApp 스토리지 및 IP 네트워크에 연결된 UCS(Unified Computing System) 컴퓨터로 구성됩니다.



가장 일반적인 FlexPod는 FI(Fabric Interconnect)를 통해 Nexus 5500 스위치에 NetApp 파일러에 연결된 Cisco UCS B-Series 새시로 구성됩니다. FlexPod Express라는 다른 솔루션은 Nexus 3000 스위치에 연결된 UCS C-Series 새시를 사용합니다. 이 문서에서는 가장 일반적인 FlexPod에 대해 설명합니다.

성능 고려 사항

FlexPod에서 일반적으로 볼 수 있는 것처럼 여러 책임 당사자가 있는 복잡한 환경에서는 문제를 해결하려면 여러 측면을 고려해야 합니다.레이어 2 및 IP 네트워크의 일반적인 성능 문제는 다음과 같습니다.

- 패킷 또는 프레임 손실 - 데이터 비트가 손실되면 애플리케이션 성능이 저하됩니다.
- 버퍼링 - 패킷 또는 프레임이 큐/버퍼에 너무 많은 시간을 보내는 경우 특히 스토리지 네트워킹의 경우 애플리케이션에서 특정 성능 영향을 볼 수 있습니다.레이턴시, 순서 조정 및 노멀라이저 문제는 이 범주에 속합니다.
- MTU 불일치 문제 및 단편화 - 더 높은 성능에 도달할 때 흔히 발생하는 문제입니다.단편화 및 MTU 불일치와 관련된 문제는 이 범주에 속합니다.

환경

성능을 측정할 환경을 파악하는 것이 중요합니다.스토리지 유형 및 프로토콜은 물론 영향을 받는 서버의 OS(운영 체제) 및 위치에 대한 질문을 제기하여 문제를 올바르게 해결해야 합니다.연결을 보여주는 토폴로지 다이어그램은 최소 수준입니다.

측정

측정되는 것과 측정되는 방법을 알아야 합니다.대부분의 스토리지 및 하이퍼바이저 공급업체는 물론 특정 애플리케이션도 시스템의 성능/상태를 나타내는 어떤 종류의 측정을 제공합니다.이러한 측정은 대부분의 트러블슈팅 방법론을 대체하지 않으므로 시작하기에 적합한 지점입니다.

예를 들어 하이퍼바이저에서 NFS(Network File System) 스토리지 레이턴시 측정은 성능이 저하될 수 있지만 자체적으로 네트워크를 구현하지는 않습니다.NFS의 경우 호스트에서 NFS 스토리지 IP 네트워크로 간단한 ping을 수행하면 네트워크가 잘못되었는지 여부를 나타낼 수 있습니다.

초기 계획

특히 TAC 케이스를 열 때 이 점을 충분히 강조할 수는 없습니다.성능이 만족스럽지 않음을 나타내려면 측정된 매개변수를 표시해야 합니다.여기에는 예상 및 테스트된 값이 포함됩니다.이전 데이터와 해당 데이터를 달성하는 데 사용된 테스트 방법론을 보여주는 것이 좋습니다.

예를 들면 단일 이니시에이터에서 단일 LUN(Logical Unit Number)까지 쓰기 전용을 사용하여 테스트한 경우 10ms의 레이턴시가 충족되면 완전히 로드된 시스템에 대해 레이턴시가 예상된다는 것을 나타낼 수 없습니다.

FlexPod의 성능 문제

이 문서는 대부분의 FlexPod 환경을 대상으로 작성되었으므로 데이터 센터 솔루션을 담당하는 TAC 팀에서 가장 자주 발생하는 문제만 간략하게 설명합니다.

일반적인 문제

이 섹션에서는 스토리지 및 IP/레이어 2 네트워크에 공통적으로 적용되는 문제에 대해 설명합니다.

프레임 및 패킷 손실

프레임 및 패킷 손실은 성능에 영향을 주는 가장 빈번한 요소입니다. 일반적으로 문제 지표를 찾는 위치 중 하나는 인터페이스 레벨입니다. Nexus 5000 또는 UCS Nexus 운영 체제(NX-OS) CLI에서 **show interface**를 입력합니다 | 초 "작동 중" | **egrep ^((Eth|fc)|discard|drop|CRC** 명령 작동 중인 인터페이스의 경우 이름을 나열하고 카운터와 삭제를 삭제합니다. 마찬가지로 모든 인터페이스에 대한 오류 통계를 표시하는 **show interface counters error** 명령을 입력하면 개요가 표시됩니다.

이더넷 세계

0이 아닌 카운터가 문제를 나타내지 않을 수 있다는 점을 알아야 합니다. 일부 시나리오에서는 이러한 카운터가 초기 설정 또는 이전 운영 변경 시 제거되었을 수 있습니다. 카운터를 늘려야 합니다.

또한 ASIC 레벨에서 카운터를 수집할 수 있으며, 이는 더욱 암시적일 수 있습니다. 특히 인터페이스의 CRC(Cyclic Redundancy Check) 오류의 경우, 입력할 TAC 즐겨찾기 명령은 **show hardware internal carmel crc**입니다. Carmel은 포트 레벨 전달을 담당하는 ASIC의 이름입니다.

포트 단위로 6100 Series FI 또는 Nexus 5600 스위치에서 유사한 출력을 가져올 수 있습니다. FI 6100, gatos ASIC의 경우 다음 명령을 입력합니다.

```
show hardware internal gatos port ethernet X/Y | grep
"OVERSIZE|TOOLONG|DISCARD|UNDERSIZE|FRAGMENT|T_CRC|ERR|JABBER|PAUSE"
```

Nexus 5600의 경우 bigsur ASIC에서 다음 명령을 입력합니다.

```
show hardware internal bigsur port eth x/y | egrep
"OVERSIZE|TOOLONG|DISCARD|UNDERSIZE|FRAGMENT|T_CRC|ERR|JABBER|PAUSE"
```

Carmel ASIC에 대한 명령은 CRC 패킷이 수신된 위치, 패킷이 전달된 위치, 특히 고객이 받을 들었는지 여부를 표시합니다.

Nexus 5000 및 UCS NX-OS 작업이 모두 컷스루(cut-through)되므로 잘못된 FCS(Frame Check Sequence)가 있는 스위칭 모드 프레임은 포워딩 전에만 맞춤화됩니다. 손상된 프레임의 출처를 찾는 것이 중요합니다.

```
bdsol-6248-06-A(nxos)# show hardware internal carmel crc
```

```
+-----+-----+-----+-----+-----+-----+-----+-----+
| Port   | MM rx CRC | MM Rx Stomp| FI rx CRC | FI Rx Stomp | FI tx CRC | FI tx Stomp| MM tx CRC |
+-----+-----+-----+-----+-----+-----+-----+-----+
(....)
| Eth 1/17 |    --- |    --- |    --- |    908100 |    --- |    --- |    --- |
| Eth 1/18 |    --- |    --- |    --- |    298658 |    --- |    --- |    --- |
(....)
| Eth 1/34 |    --- |    --- |    --- |    --- |    --- | 1206758 | 1206758 |
```

이 예에서는 Nexus 5000에 대한 업링크인 Eth 1/17 및 Eth 1/18에서 오는 맞춤화된 패킷을 보여줍니다. 이러한 프레임이 나중에 Eth 1/17 + Eth 1/18 rx Stomp = Eth 1/34 Tx Stomp와 같이 Eth 1/34로 전송되었다고 가정할 수 있습니다.

Nexus 5000에서도 이와 유사한 현상이 나타났습니다.

```
bdsol-n5548-05# show hardware internal carmel crc
```

Port	MM rx CRC	MM Rx Stomp	FI rx CRC	FI Rx Stomp	FI tx CRC	FI tx Stomp	MM tx CRC
(....)							
Eth 1/14	13	---	---	13	---	---	---
(.....)							
Eth 1/19	7578	---	---	7463	---	---	---

이 출력은 두 개의 링크에서 수신되고 전달 전에 고객으로 표시된 CRC를 표시합니다. 자세한 내용은 [Nexus 5000 문제 해결 가이드를 참조하십시오](#).

파이버 채널 월드

삭제(discrds, error, CRCs, B2B 신용 소모)를 찾는 간단한 방법은 `show interface counters fc` 명령을 통해 확인할 수 있습니다.

Nexus 5000 및 Fabric Interconnect에서 사용할 수 있는 이 명령은 파이버 채널 세계에서 발생하는 상황을 잘 보여줍니다.

예:

```
bdsol-n5548-05# show interface counters fc | i fc|disc|error|B2B|rate|put
fc2/16
1 minute input rate 72648 bits/sec, 9081 bytes/sec, 6 frames/sec
1 minute output rate 74624 bits/sec, 9328 bytes/sec, 5 frames/sec
96879643 frames input, 155712103332 bytes
0 discards, 0 errors, 0 CRC
113265534 frames output, 201553309480 bytes
0 discards, 0 errors
0 input OLS, 1 LRR, 0 NOS, 0 loop inits
1 output OLS, 2 LRR, 0 NOS, 0 loop inits
0 transmit B2B credit transitions from zero
0 receive B2B credit transitions from zero
16 receive B2B credit remaining
32 transmit B2B credit remaining
0 low priority transmit B2B credit remaining
(...)
```

이 인터페이스는 사용 중이 아니며, 출력에서 폐기 또는 오류가 발생하지 않았음을 표시합니다.

또한 0의 B2B 신용 전환이 강조되었습니다. Cisco 버그 ID CSCue80063 및 CSCut08353으로 인해 해당 카운터를 신뢰할 수 없습니다. 이 카운터는 Cisco MDS에서 작동하지만 Nexus5k 플랫폼의 UCS에서는 작동하지 않습니다. 또한 Cisco 버그 ID CSCsz95889를 확인할 수 있습니다.

파이버 채널(FC)을 위한 이더넷 세계의 카멜과 마찬가지로 fc-mac 시설을 사용할 수 있습니다. 예를 들어 포트 fc2/1의 경우 `show hardware internal fc-mac 2 port 1 statistics` 명령을 입력합니다. 표시되는 카운터는 16진수 형식입니다.

```
bdsol-6248-06-A(nxos)# show interface fc1/32 | i disc
```

```

15 discards, 0 errors
0 discards, 0 errors
bdsol-6248-06-A(nxos)# show hardware internal fc-mac 1 port 32 statistics
ADDRESS          STAT                                          COUNT

```

```

0x0000003d FCP_CNTR_MAC_RX_BAD_WORDS_FROM_DECODER          0x70
0x00000042 FCP_CNTR_MAC_CREDIT_IG_XG_MUX_SEND_RRDY_REQ 0x1e4f1026
0x00000043 FCP_CNTR_MAC_CREDIT_EG_DEC_RRDY          0x66cafd1
0x00000061 FCP_CNTR_MAC_DATA_RX_CLASS3_FRAMES        0x1e4f1026
0x00000069 FCP_CNTR_MAC_DATA_RX_CLASS3_WORDS      0xe80946c708
0x000d834c FCP_CNTR_PIF_RX_DROP                          0xf
0x00000065 FCP_CNTR_MAC_DATA_TX_CLASS3_FRAMES        0x66cafd1
0x0000006d FCP_CNTR_MAC_DATA_TX_CLASS3_WORDS      0x2b0fae9588
0xffffffff FCP_CNTR_OLS_IN                          0x1
0xffffffff FCP_CNTR_LRR_IN                          0x1
0xffffffff FCP_CNTR_OLS_OUT                        0x1

```

출력에 입력 시 15개의 폐기 결과가 표시됩니다. 이 값은 0xf(15진수)로 계산된 FCP_CNTR_PIF_RX_DROP과 일치할 수 있습니다. 이 정보는 FWM(Forwarding Manager) 정보와 다시 상호 연결될 수 있습니다.

```

bdsol-6248-06-A(nxos)# show platform fwm info pif fc 1/32 verbose | i drop|discard|asic
fc1/32 pd: slot 0 logical port num 31 slot_asic_num 3 global_asic_num 3 fwm_inst 7
fc 0
fc1/32 pd: tx stats: bytes 191196731188 frames 107908990 discard 0 drop 0
fc1/32 pd: rx stats: bytes 998251154572 frames 509332733 discard 0 drop 15
fc1/32 pd fcoe: tx stats: bytes 191196731188 frames 107908990 discard 0 drop 0
fc1/32 pd fcoe: rx stats: bytes 998251154572 frames 509332733 discard 0 drop 15

```

그러나 이는 관리자에게 삭제 양과 해당 ASIC 번호를 알려줍니다. 삭제된 ASIC의 원인에 대한 가져 오기 정보를 쿼리해야 합니다.

```

bdsol-6248-06-A(nxos)# show platform fwm info ASIC-errors 3
Printing non zero Carmel error registers:
DROP_SHOULD_HAVE_INT_MULTICAST: res0 = 25 res1 = 0 [36]
DROP_INGRESS_ACL: res0 = 15 res1 = 0 [46]

```

이 경우 트래픽은 인그레스 ACL(Access Control List)에 의해 삭제되었으며, 일반적으로 FC 환경에서는 zoning입니다.

MTU 불일치

FlexPod 환경에서는 필요한 애플리케이션 및 프로토콜에 대한 엔드 투 엔드 MTU(Maximum Transition Unit) 설정을 수용하는 것이 중요합니다. 대부분의 환경에서는 FCoE(Fibre Channel over Ethernet) 및 점보 프레임입니다.

또한 단편화가 발생할 경우 성능이 저하될 것으로 예상됩니다. NFS(Network File System) 및 iSCSI(Internet Small Computer System Interface)와 같은 프로토콜의 경우 엔드 투 엔드 IP MTU(Maximum Transmission Unit) 및 TCP MSS(Maximum Segment Size)를 테스트하고 검증하는 것이 중요합니다.

점보 프레임 또는 FCoE의 문제를 해결하든, 제대로 작동하려면 환경 전반에 걸쳐 일관된 구성 및 CoS(Class of Service) 표시가 필요하다는 점을 기억해야 합니다.

UCS와 Nexus의 경우 인터페이스별로 유효성을 검사하는 데 유용한 명령이며 QoS 그룹 MTU당 설정은 `show queuing interface`입니다 |i queuing|qos-group|MTU.

Nexus 5000 및 UCS 플랫폼의 MTU 디스플레이

UCS와 Nexus의 알려진 측면은 인터페이스에 MTU를 표시하는 것입니다. 이 출력은 점보 프레임 및 FCoE를 대기하도록 구성된 인터페이스를 보여줍니다.

```
bdsol-6248-06-A(nxos)# show queuing interface e1/1 | i MTU
q-size: 360640, HW MTU: 9126 (9126 configured)
q-size: 79360, HW MTU: 2158 (2158 configured)
```

동시에 **show interface** 명령은 1,500바이트를 표시합니다.

```
bdsol-6248-06-A(nxos)# show int e1/1 | i MTU
MTU 1500 bytes, BW 10000000 Kbit, DLY 10 usec
```

ASIC는 Carmel ASIC 정보와 비교할 경우 지정된 포트의 MTU 기능을 표시합니다.

```
show hardware internal carmel port ethernet 1/1 | egrep -i MTU
mtu : 9260
```

이 MTU 불일치는 앞서 언급한 플랫폼에서 예상되며, 잠재적으로 네이핑을 잘못 인식할 수 있습니다.

엔드 투 엔드 구성

엔드 투 엔드 일관된 컨피그레이션은 적절한 성능을 보장하는 유일한 방법입니다. Cisco 측 및 VMware ESXi에 대한 점보 프레임 컨피그레이션 및 단계는 VMware ESXi [엔드 투 엔드 점보 MTU 컨피그레이션 예에 설명되어 있습니다](#).

[UCS FCoE 업링크 컨피그레이션의 예](#)는 UCS 및 Nexus 5000 컨피그레이션입니다. 기본 Nexus 5000 컨피그레이션의 개요는 참조된 문서의 부록 A를 참조하십시오.

[Cisco UCS Blade에 대한 FCoE 연결 설정은 FCoE를 위한 UCS 컨피그레이션에 중점을 둡니다](#). [.Nexus 5000 NPIV FCoE with FCoE FCoE FCoE Attached UCS Configuration 예제](#)는 Nexus 컨피그레이션에 중점을 둡니다.

엔드 투 엔드 점보 프레임 테스트

최신 운영 체제는 간단한 ICMP(Internet Control Message Protocol) 테스트를 통해 적절한 점보 프레임 구성을 테스트할 수 있습니다.

계산

9000바이트 - 옵션이 없는 IP 헤더(20바이트) - ICMP 헤더(8바이트) = 데이터 8972바이트

인기 운영 체제의 명령

리눅스

```
ping a.b.c.d -M do -s 8972
```

Microsoft Windows

```
ping -f -l 8972 a.b.c.d
```

ESXi

```
vmkping -d -s 8972 a.b.c.d
```

버퍼 관련 문제

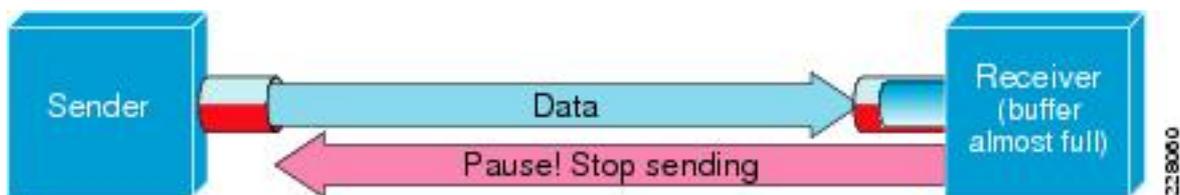
버퍼링 및 기타 레이턴시 관련 문제는 FlexPod 환경에서 발생하는 일반적인 성능 저하의 원인입니다. 레이턴시가 실제 버퍼링 문제 때문에 보고되는 일부 문제는 아니며, 상당한 측정은 엔드 투 엔드 레이턴시를 나타낼 수 있습니다. 예를 들어 NFS의 경우 실제 네트워크 레이턴시가 아닌 스토리지에 대한 읽기/쓰기를 성공적으로 수행하려면 보고된 기간이 필요할 수 있습니다.

혼잡은 버퍼링의 가장 일반적인 원인입니다. 레이어 2 환경에서는 혼잡 때문에 프레임이 버퍼링되고 심지어 꼬리가 떨어질 수 있습니다. 혼잡 기간 동안 삭제를 방지하기 위해 IEEE 802.3x 일시 중지 프레임 및 PFC(Priority Flow Control)가 도입되었습니다. 양쪽 모두 혼잡이 지속되는 동안 단시간 동안 전송을 보류하도록 중단점에 요청해야 합니다. 이는 네트워크 혼잡(수신한 데이터를 데이터 양에 압도함) 또는 FCoE의 경우와 같이 우선 순위가 지정된 프레임을 통과해야 하기 때문에 발생할 수 있습니다.

플로우 제어 - 802.3x

어떤 인터페이스에서 흐름 제어를 활성화했는지 확인하려면 **show interface flowcontrol** 명령을 입력합니다. 플로우 제어의 활성화 여부는 스토리지 공급업체의 권장 사항을 따르는 것이 중요합니다.

여기에서는 802.3x 흐름 제어의 작동 방식을 보여 주는 그림입니다.

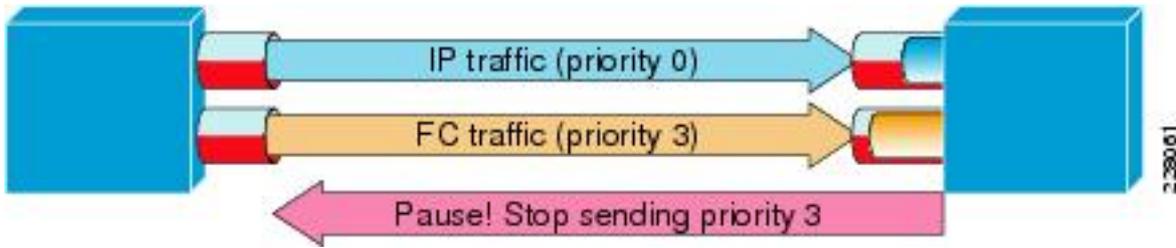


PFC - 802.1Qbb

PFC는 모든 설정에 필요하지 않지만 대부분의 경우 권장됩니다. PFC가 활성화된 인터페이스를 확인하려면 **show interface priority-flow-control | i On** 명령은 UCS의 NX-OS 및 Nexus 5000에서 실행할 수 있습니다.

FI와 Nexus 5000 간의 인터페이스는 해당 목록에 표시됩니다. 그렇지 않은 경우 QoS 컨피그레이션을 확인해야 합니다. PFC를 활용하려면 QoS가 일관성 있는 엔드 투 엔드 방식을 갖춰야 합니다. PFC가 특정 인터페이스에 나타나지 않는 이유를 확인하려면 **show system internal dcbx log interface ethernet x/y** 명령을 입력하여 DCBX(Data Center Bridging Capabilities Exchange Protocol) 로그를 가져옵니다.

여기에 PFC를 사용하여 일시 중지 프레임이 작동하는 방법을 보여 주는 그림이 표시됩니다.



`show interface priority-flow-control` 명령을 사용하면 관리자가 우선 순위 일시 중지 프레임의 QoS 클래스별 동작을 관찰할 수 있습니다.

예를 들면 다음과 같습니다.

```
bdsol-6120-05-A(nxos)# show queuing interface ethernet 1/1 | i prio
Per-priority-pause status : Rx (Inactive), Tx (Inactive)
Per-priority-pause status : Rx (Inactive), Tx (Active)
```

이 출력은 디바이스가 PPP 프레임 전송(TX)하는 것을 2차 클래스에서도 보여줍니다.

이 경우 이더넷 1/1은 IOM을 향하는 포트이며, 전체 포트는 PFC를 활성화하지 않지만 FEX 포트에 대해 PPP 프레임을 처리할 수 있습니다.

```
bdsol-6120-05-A(nxos)# show interface e1/1 priority-flow-control
=====
Port Mode Oper(VL bmap) RxPPP TxPPP
=====
Ethernet1/1 Auto Off 4885 3709920
```

이 경우 FEX 인터페이스가 포함됩니다.

```
bdsol-6120-05-A(nxos)# show interface priority-flow-control | egrep .*\/.*\/
Ethernet1/1/1 Auto Off 0 0
Ethernet1/1/2 Auto Off 0 0
Ethernet1/1/3 Auto Off 0 0
Ethernet1/1/4 Auto Off 0 0
Ethernet1/1/5 Auto On (8) 8202210 15038419
Ethernet1/1/6 Auto On (8) 0 1073455
Ethernet1/1/7 Auto Off 0 0
Ethernet1/1/8 Auto On (8) 0 3956077
Ethernet1/1/9 Auto Off 0 0
```

관련된 FEX 포트는 `show fex X detail`을 통해 확인할 수도 있습니다. 여기서 X는 쉐시 번호입니다.

```
bdsol-6120-05-A(nxos)# show fex 1 detail | section "Fex Port"
Fex Port State Fabric Port
Eth1/1/1 Down Eth1/1
Eth1/1/2 Down Eth1/2
Eth1/1/3 Down None
Eth1/1/4 Down None
Eth1/1/5 Up Eth1/1
Eth1/1/6 Up Eth1/2
Eth1/1/7 Down None
Eth1/1/8 Up Eth1/2
Eth1/1/9 Up Eth1/2
```

일시 중지 메커니즘에 대한 자세한 내용은 다음 문서를 참조하십시오.

- [Fibre Channel over Ethernet 운영](#)
- [유니파이드 패브릭 백서-FCoE\(Fibre Channel over Ethernet\)](#)

대기 중 폐기

Nexus 5000과 UCS NX-OS 모두 QOS 그룹별로 대기하면서 인그레스(ingress) 폐기를 지속적으로 추적합니다.예:

```

bdsol-6120-05-A(nxos)# show queuing interface
Ethernet1/1 queuing information:
TX Queuing
  qos-group  sched-type  oper-bandwidth
    0          WRR        50
    1          WRR        50
RX Queuing
  qos-group 0
  q-size: 243200, HW MTU: 9280 (9216 configured)
  drop-type: drop, xon: 0, xoff: 243200
Statistics:
  Pkts received over the port          : 31051574
  Ucast pkts sent to the cross-bar     : 30272680
  Mcast pkts sent to the cross-bar     : 778894
  Ucast pkts received from the cross-bar : 27988565
  Pkts sent to the port                : 34600961
  Pkts discarded on ingress          : 0
  Per-priority-pause status           : Rx (Inactive), Tx (Active)

```

인그레스 삭제는 삭제를 허용하도록 구성된 대기열에서만 발생해야 합니다.

인그레스 대기열 삭제는 다음 이유로 인해 발생할 수 있습니다.

- 일부 인터페이스에서 활성화된 SPAN(Switched Port Analyzer)/모니터링 세션(Cisco 버그 ID [CSCur25521 참조](#))
- 다른 인터페이스의 뒷면 압력으로, 일시 중지 프레임은 일반적으로 활성화될 때 표시됩니다.
- CPU로 전송된 트래픽

드라이버 문제

Cisco는 UCS, Nic 및 fnic를 위한 두 가지 운영 체제 드라이버를 제공합니다.enic은 이더넷 연결을 담당하고 fnic는 파이버 채널 및 FCoE 연결을 담당합니다.UCS [상호 운용성 매트릭스](#)에 명시되어 있는 것과 정확히 동일한 엔틱 및 nic 드라이버가 매우 중요합니다.잘못된 드라이버로 인해 발생하는 문제는 패킷 손실 및 레이턴시 추가에서 더 긴 부팅 프로세스 또는 완전한 연결 부족까지 다양합니다.

어댑터 정보

Cisco 제공 어댑터는 전달되는 트래픽과 드롭에 대한 좋은 측정치를 제공할 수 있습니다.이 예에서는 새시 X, 서버 Y 및 어댑터 Z에 연결하는 방법을 보여줍니다.

```

bdsol-6248-06-A# connect adapter X/Y/Z
adapter X/Y/Z # connect
No entry for terminal type "dumb";

```

using dumb terminal settings.

여기에서 관리자는 MCP(Monitoring Center for Performance) 기능에 로그인할 수 있습니다.

```
adapter 1/2/1 (top):1# attach-mcp
No entry for terminal type "dumb";
using dumb terminal settings
```

MCP 기능을 사용하면 LIF(논리적 인터페이스)당 트래픽 사용량을 모니터링할 수 있습니다.

```
adapter 1/2/1 (mcp):1# vnic
(...)
```

```
-----
```

v n i c			l i f			v i f					
id	name	type	bb:dd.f	state	lif	state	uif	ucsm	idx	vlan	state
13	vnic_1	enet	06:00.0	UP	2	UP	=>0	834	20	3709	UP
14	vnic_2	fc	07:00.0	UP	3	UP	=>0	836	17	970	UP

```
-----
```

샐시 1, 서버 1 및 어댑터 1에는 가상 인터페이스(가상 이더넷 또는 가상 파이버 채널) 834 및 836과 연결된 2개의 VNIC(Virtual Network Interface Card)가 있습니다. 이러한 카드는 번호 2와 3입니다. LIF 2 및 3에 대한 통계는 다음과 같이 확인할 수 있습니다.

```
adapter 1/2/1 (mcp):3# lifstats 2
```

DELTA	TOTAL	DESCRIPTION
4	4	Tx unicast frames without error
53999	53999	Tx multicast frames without error
69489	69489	Tx broadcast frames without error
500	500	Tx unicast bytes without error
8361780	8361780	Tx multicast bytes without error
22309578	22309578	Tx broadcast bytes without error
2	2	Rx unicast frames without error
2791371	2791371	Rx multicast frames without error
4595548	4595548	Rx broadcast frames without error
188	188	Rx unicast bytes without error
260068999	260068999	Rx multicast bytes without error
514082967	514082967	Rx broadcast bytes without error
3668331	3668331	Rx frames len == 64
2485417	2485417	Rx frames 64 < len <= 127
655185	655185	Rx frames 128 <= len <= 255
434424	434424	Rx frames 256 <= len <= 511
143564	143564	Rx frames 512 <= len <= 1023
94.599bps		Tx rate
2.631kbps		Rx rate

UCS의 관리자는 총 및 델타(두 번의 리프통계 실행 사이) 열, 현재 LIF당 트래픽 로드 및 발생한 오류에 대한 정보를 제공한다는 점에 유의해야 합니다.

앞의 예에서는 로드가 매우 적은 오류가 없는 인터페이스를 보여 줍니다. 이 예에서는 다른 서버를 보여 줍니다.

```
adapter 4/4/1 (mcp):2# lifstats 2
```

DELTA	TOTAL	DESCRIPTION
127927993	127927993	Tx unicast frames without error
273955	273955	Tx multicast frames without error
122540	122540	Tx broadcast frames without error
50648286058	50648286058	Tx unicast bytes without error
40207322	40207322	Tx multicast bytes without error
13984837	13984837	Tx broadcast bytes without error

```

28008032          28008032 Tx TSO frames
262357491        262357491 Rx unicast frames without error
55256866         55256866 Rx multicast frames without error
51088959         51088959 Rx broadcast frames without error
286578757623    286578757623 Rx unicast bytes without error
4998435976      4998435976 Rx multicast bytes without error
7657961343      7657961343 Rx broadcast bytes without error

```

96 96 Rx rq drop pkts (no bufs or rq disabled)

```

136256          136256 Rx rq drop bytes (no bufs or rq disabled)
5245223        5245223 Rx frames len == 64
136998234      136998234 Rx frames 64 < len <= 127
9787080        9787080 Rx frames 128 <= len <= 255
14176908       14176908 Rx frames 256 <= len <= 511
11318174       11318174 Rx frames 512 <= len <= 1023
61181991       61181991 Rx frames 1024 <= len <= 1518
129995706      129995706 Rx frames len > 1518

```

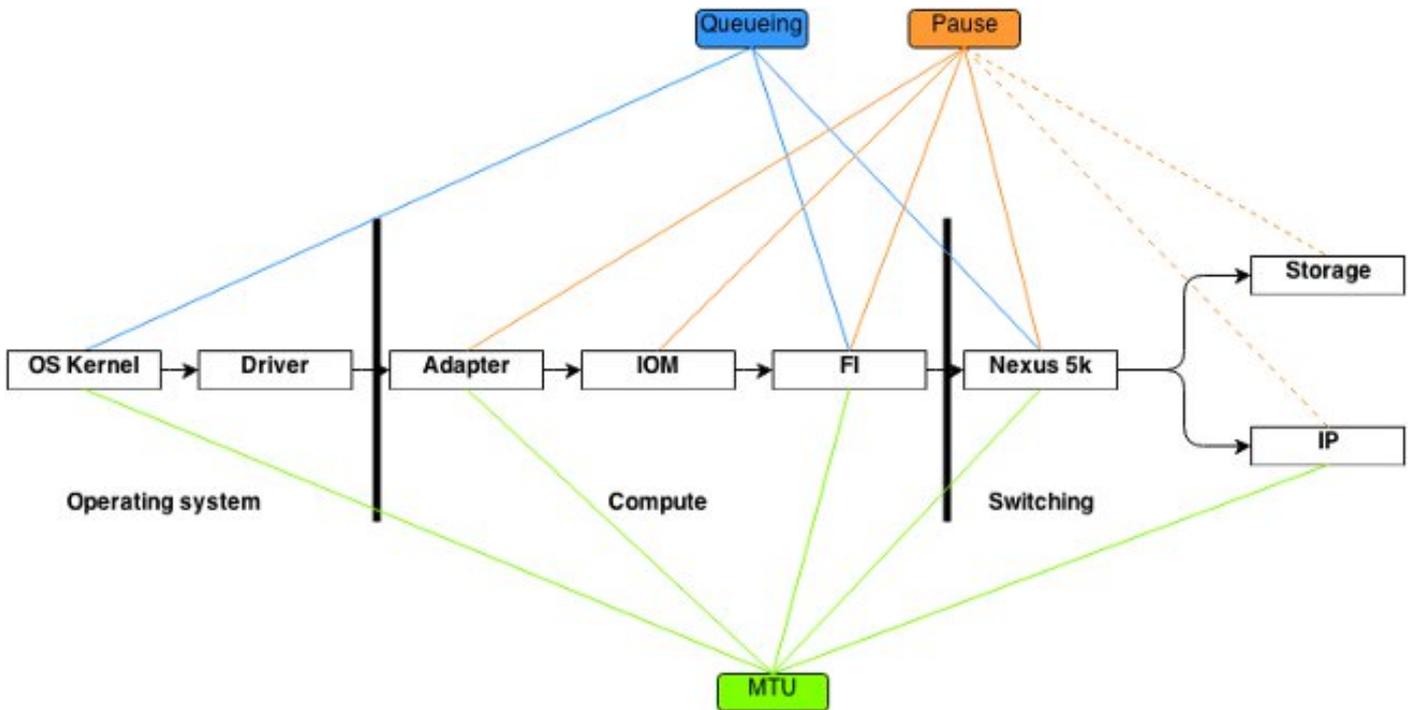
136.241kbps Tx rate

784.185kbps Rx rate

두 가지 흥미로운 정보 비트에 따르면, 96개의 프레임이 버퍼링 또는 버퍼링 기능이 비활성화되어 있고 추가로 처리되는 TCP 세그먼트 오프로드(TSO) 세그먼트로 인해 어댑터에 의해 삭제되었습니다.

논리적 패킷 흐름

여기에 표시된 다이어그램은 FlexPod 환경의 논리적 패킷 흐름을 보여줍니다.



이 다이어그램은 FlexPod 환경을 통해 프레임이 통과하는 구성 요소의 분석을 나타냅니다. 이는 블록의 복잡성을 반영하지 않으며 특정 기능을 구성 및 검증해야 하는 위치를 암기하는 단순한 방법입니다.

입력/출력 모듈

- 트래픽 손실 - `show platform software [woodside|redwood] loss` 명령을 입력합니다. 이 명령을 실행하면 손실 카운터가 0으로 설정됩니다. 이를 통해 인터페이스별로 일시 중지 프레임 및 삭제제를 확인할 수 있습니다

```

# show platform software woodside loss

```

Port	S3803			Input Extra Drop	S3 Loss Counters	Loss											
	Tx Pkts	Rx Pkts	Errors			Counters	COB										
						0	1	2	3	4	5	6	7	8	9	10	11
0-M2	0	82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0-HI23	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

기본 인프라의 작동 방식 때문에 두 명령 실행 사이에 손실이 발생한 인터페이스에 대해서만 카운터가 표시됩니다. 이 예에서는 NI2 인터페이스가 82개의 일시 중지 프레임을 수신했고 28개의 일시 중지 프레임이 HI23 인터페이스로 전송되었으며, 이는 블레이드 3에 연결되어 있음을 알 수 있습니다.

설계 고려 사항

FlexPod를 사용하면 유연한 구성과 스토리지 및 데이터 네트워킹을 설정할 수 있습니다. 유연성과 더불어 추가적인 과제가 발생합니다. 모범 사례 문서와 CVD(Cisco Validated Design)를 따르는 것이 중요합니다.

- CVD - [FlexPod 구축 설명서](#)
- NetApp 스토리지 모범 사례(Flexpod에 국한되지 않음) - [Cisco UCS\(Unified Computing System\) 스토리지 연결 옵션 및 NetApp 스토리지 모범 사례](#)

포트 속도 선택 및 포트 채널 고려 사항

TAC 엔지니어가 흔히 볼 수 있는 문제는 모범 사례 문서에서 참조되는 10Gbit 이더넷 대신 1Gbit 이더넷을 선택하여 링크를 과도하게 사용하는 것입니다. 예를 들어, 10Gbit 링크 1개와 비교하여 10개의 1Gbit 링크에서는 단일 플로우 성능이 더 우수하지 않습니다. 포트 채널에서 단일 플로우는 단일 링크를 통해 이동할 수 있습니다.

Nexus 및/또는 FI의 NX-OS에서 어떤 로드 밸런싱 방법이 사용되는지 확인하려면 `show port-channel load-balance` 명령을 입력합니다. 관리자는 패킷 또는 프레임의 발신 인터페이스로 선택할 포트 채널의 인터페이스를 확인할 수도 있습니다. 두 호스트 간의 VLAN49에 있는 프레임의 간단한 예는 다음과 같습니다.

```

show port-channel load-balance forwarding-path interface port-channel 928 vlan 49
src-mac 70ca.9bce.ee24 dst-mac 8478.ac55.2fc2
Missing params will be substituted by 0's.
Load-balance Algorithm on switch: source-dest-ip
crc8_hash: 2      Outgoing port id: Ethernet1/27
Param(s) used to calculate load-balance:
dst-mac: 8478.ac55.2fc2

```

스토리지 관련 문제

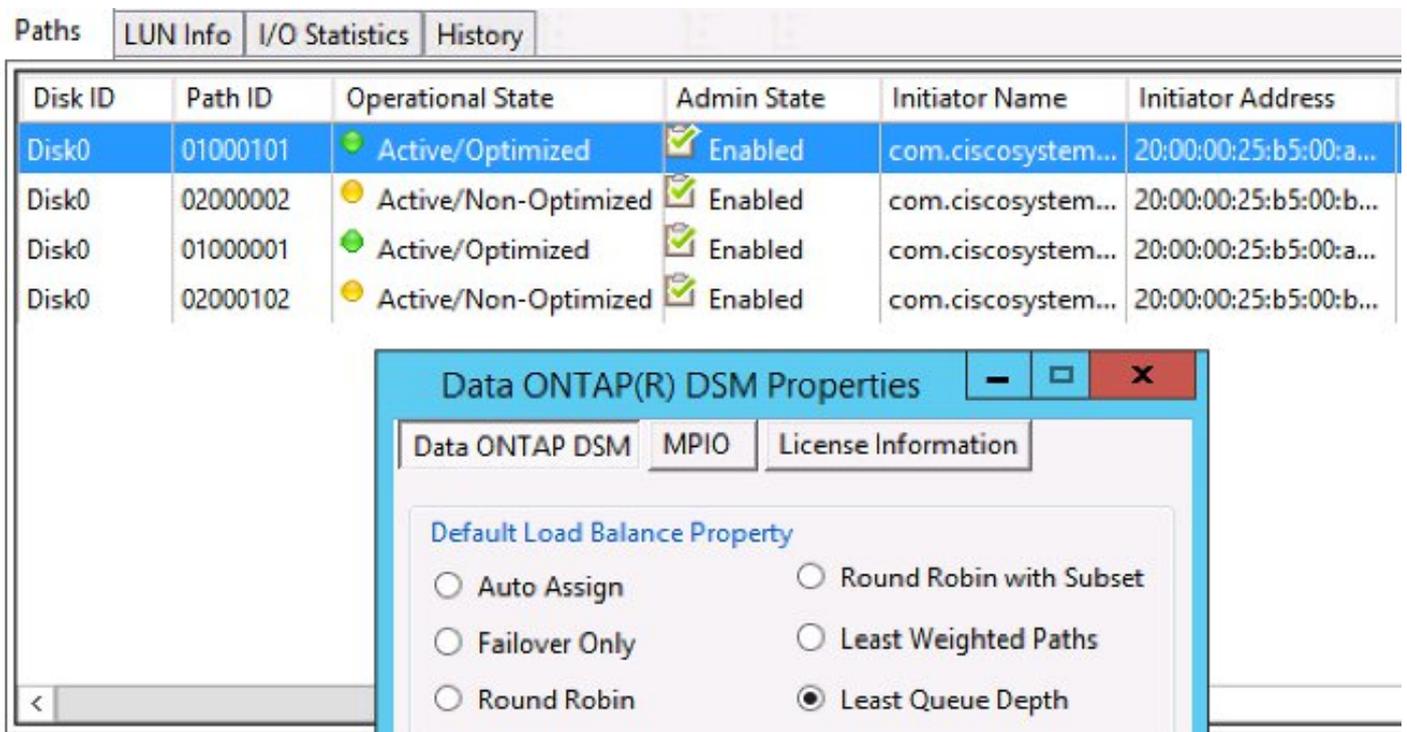
앞서 설명한 문제는 데이터 및 스토리지 네트워킹 모두에서 흔히 발생합니다.완전성을 위해 SAN(Storage Area Network)과 관련된 성능 문제도 언급되었습니다.스토리지 프로토콜은 복원력 및 다중 경로 지정 기능을 통해 구축되었습니다.ALUA(Asymmetric Logical Unit Assignment) 및 MPIO(Multi-path IO) 등의 기술이 등장함에 따라 관리자에게 더 많은 유연성과 옵션이 제공됩니다.

스토리지 배치

또 다른 고려 사항은 스토리지를 배치하는 것입니다.FlexPod 설계에서는 스토리지를 Nexus 스위치에 연결해야 합니다.직접 연결된 스토리지는 CVD를 준수하지 않습니다.모범 사례를 따르는 경우, 직접 연결된 스토리지가 있는 설계가 지원됩니다.이와 동시에 이러한 설계는 FlexPod가 아닙니다.

최적 경로 선택

대부분의 옵션은 Cisco 장치에 영향을 미치지 않으므로, 기술적으로 Cisco 문제가 아닙니다.최적의 경로를 선택하고 유지하는 것은 일반적인 문제입니다.최신 DSM(Device Specific Module)은 다중 경로를 제공할 수 있으며 복원력 및 로드 밸런싱을 제공하기 위해 특정 기준에 따라 최적의 경로를 선택해야 합니다.이 스크린샷은 NetApp DSM for Microsoft Windows 및 로드 밸런싱 옵션에 사용할 수 있는 네 가지 경로를 보여줍니다.



권장 설정은 스토리지 공급업체와의 논의를 기반으로 선택해야 합니다.이러한 설정은 성능 문제에 영향을 미칠 수 있습니다.TAC에서 수행하도록 요청할 수 있는 일반적인 테스트는 패브릭 A 또는 패브릭 B를 통한 읽기/쓰기 테스트입니다. 이 테스트에서는 일반적으로 성능 문제를 이 문서의 "공통 문제" 섹션에서 설명하는 상황으로 좁힐 수 있습니다.

VM 및 하이퍼바이저 트래픽 공유

이 점은 벤더에 관계없이 컴퓨팅 구성 요소에만 적용됩니다.컴퓨팅 관점에서 하이퍼바이저에 대한 스토리지 네트워크를 쉽게 설정할 수 있는 방법은 각 파이버에 하나씩 두 개의 HBA(Host Bus Adapter)를 생성하고 이 두 인터페이스를 통해 부팅 LUN 트래픽과 가상 머신(VM) 스토리지 트래픽을 모두 실행하는 것입니다.부팅 LUN 트래픽 및 VM 스토리지 트래픽을 분할하는 것이 좋습니다.따라서 성능이 향상되고 두 종류의 트래픽 간에 논리적 분할이 가능합니다.예제는 "알려진 문제" 섹션을 참조하십시오.

문제 해결 팁

문제 축소

신속한 문제 해결과 마찬가지로, 문제를 해결하고 올바른 질문을 하는 것이 매우 중요합니다.

- 영향을 받는 디바이스/애플리케이션/VM(/아님)
- 영향을 받는 스토리지 컨트롤러(/아님)
- 영향을 받는 경로는 무엇입니까(/아님)?
- 문제가 얼마나 자주 나타납니까(/not)?

Cisco

카운터 제한 사항

이 문서 인터페이스에서는 ASIC 대기열 카운터에 대해 설명합니다.카운터는 특정 시점의 보기를 제공하므로 카운터 증가를 모니터링하는 것이 중요합니다.특정 카운터를 디자인에서 지울 수 없습니다.예를 들어, 이전에 Carmel ASIC에서 언급한 바와 같이

예를 들어, 인터페이스에 CRC 또는 폐기 기능이 있는 것은 바람직하지 않을 수 있지만, 값이 0이 아닌 것으로 예상될 수 있습니다.전환이나 초기 설정 중에 특정 시점에 카운터가 증가할 수 있습니다.따라서 카운터의 증가 및 마지막으로 카운터가 지워진 시점을 확인하는 것이 중요합니다.

컨트롤 플레인 고려 사항

카운터를 검토하는 것이 유용하지만, 특정 데이터 플레인 문제가 컨트롤 플레인 카운터와 도구를 위한 쉬운 리플렉션을 찾지 못할 수도 있다는 점을 알아야 합니다.예를 들어, Ethalyzer는 UCS와 Nexus 5000에서 모두 사용할 수 있는 매우 유용한 툴입니다.그러나 컨트롤 플레인 트래픽만 캡처할 수 있습니다.트래픽 캡처는 TAC에서 자주 요청하는 것입니다. 특히 결함이 어디에 있는지 확실하지 않은 경우 그렇습니다.

트래픽 캡처

엔드 호스트에서 탐지된 신뢰할 수 있는 트래픽 캡처는 성능 문제를 조명하고 이를 매우 빠르게 줄일 수 있습니다.Nexus 5000과 UCS 모두 트래픽 SPAN을 제공합니다.특히 UCS의 특정 HBA 및 패브릭 측 SPAN 옵션은 유용합니다.UCS에서 세션을 모니터링할 때 트래픽 캡처 기능에 대한 자세한 내용은 다음 참조를 참조하십시오.

- [물리적 및 가상 어댑터에 대한 UCS 트래픽 분석](#)(비디오)
- [Cisco UCS Manager GUI 컨피그레이션 가이드 - 트래픽 모니터링](#)

NetApp

NetApp은 스토리지 컨트롤러 문제를 해결하기 위해 다음과 같은 완벽한 유틸리티를 제공합니다.

- perfstat - 매우 유용한 유틸리티로, 일반적으로 NetApp 지원 인력용으로 실행됩니다.
- system - filer의 사용 시간 및 filer가 수행하는 작업에 대한 정보를 제공합니다. - [NetApp 지원 라이브러리](#)

가장 일반적인 명령은 다음과 같습니다.

- ```
sysstat -x 2
```

- ```
sysstat -M 2
```

sysstat -x 2 출력에서는 오버로드된 NetApp 어레이 또는 디스크를 나타낼 수 있는 몇 가지 사항을 살펴보겠습니다.

- 다음과 같은 요소가 **많은** 지속적인 CP ty 열:또는 F
- **HDD 사용률** 열이 20% 이상

이 문서에서는 NetApp을 구성하는 방법에 대해 설명합니다.[NetApp 이더넷 스토리지 모범 사례](#).

- VLAN 태깅
- VLAN 트렁킹
- 점보 MTU
- IP 해싱
- FlowControl 사용 안 함

VMware

ESXi는 문제를 해결할 수 있는 SSH(Secure Shell) 액세스를 제공합니다.관리자에게 제공되는 가장 유용한 툴은 esxtop 및 perfmon입니다.

- esxtop - Linux/BSD 상위와 마찬가지로, 사용자는 실시간 성능 관련 매개변수를 모니터링할 수 있습니다.
[esxtop을 사용하여 ESX/ESXi의 스토리지 성능 문제 파악](#)
- perfmon - 사용자가 Microsoft Windows VM(가상 머신) 문제를 해결할 수 있습니다.
[Windows Perfmon 로그 데이터를 수집하여 가상 컴퓨터 성능 문제를 진단합니다.](#)
- ESXi에서 진단 번들 수집 - vSphere [Client를 사용하여 VMware ESX/ESXi에 대한 진단 정보 수집\(653\)](#)
- Cisco B-Series 서버에 대한 VMware vSwitch 로드 밸런싱 요구 사항 - [UCS 6100 Series Fabric Interconnect를 사용하는 Cisco UCS B200 M1/M2 블레이드 서버에서는 IP 해시 기반 경로가 지원되지 않습니다.](#)

알려진 문제 및 개선 사항

- Cisco 버그 ID [CSCuj86736](#) - 패시브 twinax 케이블 CRC 오류가 증가할 수 있습니다.이는

Nexus 5000이 DFE를 최적화하지 않을 때 발생합니다."Eye height(하드웨어 내부 카멜 눈)" 매개변수가 100mv 이상인지 확인하려면 show hardware internal carmel eye 명령을 입력합니다. 이는 릴리스 5.2(1)N1(7) 및 7.0(4)N1(1)에서 수정되었습니다.

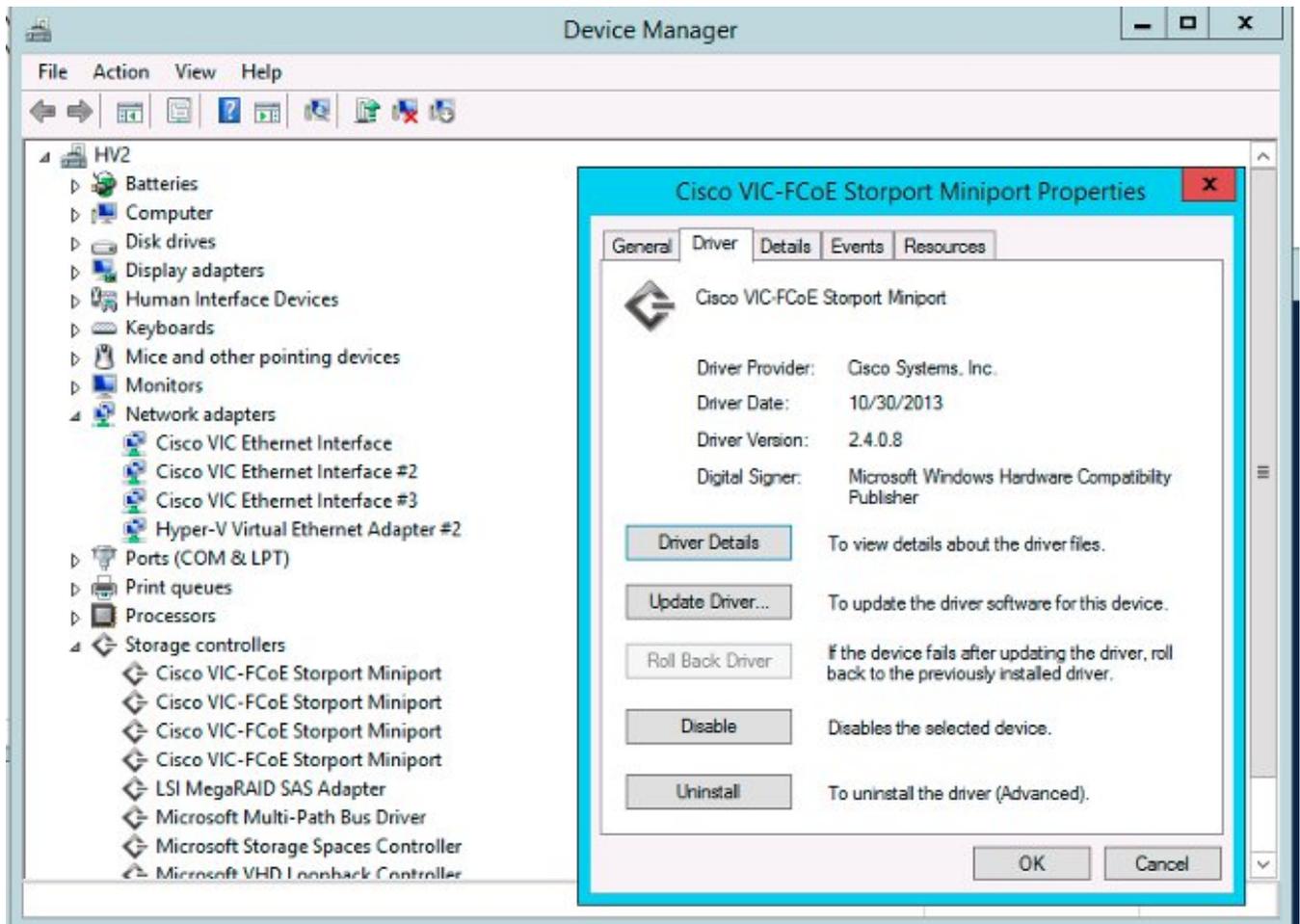
- Cisco 버그 ID [CSCUo76425](#) - 이전 버그와 유사하며 UCS 패브릭 인터커넥트에도 존재합니다. 이는 릴리스 2.2(3a)에서 수정되었습니다.
- Cisco 버그 ID [CSCUo76425](#) - 버그 [CSCUj86736](#)과 동일하며 UCS Fabric Interconnect는 예외입니다.
- Cisco 버그 ID [CSCUp40056](#) - 가상 파이버 채널 어댑터에서 [Unified Computing System 가상 머신 라이브 마이그레이션 실패](#)에 설명된 VM 트래픽과 부트 트래픽을 공유함으로써 발생하는 타이밍 문제입니다.
- 슬로우 드레인 감지 및 회피 - FC와 FCoE가 슬로우 드레인의 영향을 받는 경우가 매우 많습니다. NX-OS 릴리스 7.0(0)N1(1)에는 이를 탐지하고 방지하는 방법이 도입되었습니다. [Cisco Nexus 5500 Series NX-OS Interfaces Configuration Guide](#) 및 [Slow Drain Device Detection and Congestion Prevention](#)의 기능에 대해 자세히 [알아보십시오](#).
- Cisco 버그 ID [CSCUj81245](#) - FC 중단을 일으키는 PALO 기반 카드(VIC1240 및 기타)에 제한이 있습니다.
- Cisco 버그 ID [CSCUh61202](#) - 릴리스 2.1(3)으로 업그레이드한 후 UCS 펌웨어 FC가 중단되고 기타 여러 문제가 표시됩니다.
- Cisco 버그 ID [CSCtw91018](#) - 단일 PALO 기반 어댑터에서 VNIC에 대한 MTU 설정을 혼합하면 일부 트래픽 클래스에 문제가 발생할 수 있습니다.
- Cisco 버그 ID [CSCUq40256](#) - 패브릭 인터커넥트에서 서버 어댑터까지의 링크에서 PFC를 비활성화합니다. 이렇게 하면 스토리지 측에서 보고된 Fibre Channel 중단과 Out-of-Order 프레임으로 시작하는 다양한 문제가 발생합니다. 스토리지 연결 끊김 및 기타 성능 문제가 보고될 수 있습니다.

TAC 케이스

대부분의 경우 TAC 엔지니어가 조사를 시작하기 전에 몇 가지 기본 정보를 수집하도록 요청합니다

- 토폴로지 다이어그램 - 포트 번호 및 라인 속도를 포함하므로 반드시 필요합니다.
- UCSM 기술 지원 - [기술 지원 파일\(B 및 C 시리즈\)을 수집하기 위한 시각적 가이드](#).
- 문제가 발생하는 새시 하나에 대한 UCS 새시 기술 지원 - 이전 링크를 참조하십시오.
- Nexus 5000 기술 지원 및 UCS와 NetApp 간의 기타 네트워크 디바이스 모두 - [show tech-support details 명령의 출력을 리디렉션합니다](#).
- 두 FI 모두에서 **show queueing interface** 명령의 출력입니다.

```
connect nxos A|B
show queueing interface | no-more
show interface priority-flow-control | no-more
show interface flowcontrol | no-more.
```
- ESXi의 호스트 드라이버 버전 - 다음 명령을 입력합니다. `vmkload_mod -s enicvmkload_mod -s fnic`
- Linux - `dmesg | egrep -i 'enic|fnic'`
- Windows - "장치 관리자"에서 드라이버 버전을 확인합니다. Window 2012 R2의 예에는 Cisco VIC 이더넷 인터페이스 3개와 VIC FCoE 미니포트 인터페이스 4개(FCoE뿐만 아니라 파이버 채널 담당) 및 fnic 드라이버의 릴리스 2.4.0.8이 나와 있습니다



피드백

피드백 버튼을 사용하여 이 문서 또는 경험에 대한 피드백을 제공할 수 있습니다. Cisco는 개발 및 피드백이 수신되면 이 문서를 지속적으로 업데이트합니다.