

FlexPod общие проблемы производительности

Содержание

[Введение](#)

[FlexPod концептуальный обзор](#)

[Вопросы производительности](#)

[Среда](#)

[Измерение](#)

[Срок](#)

[Проблемы производительности в FlexPod](#)

[Типичные неполадки](#)

[Кадр и потеря пакета](#)

[Несоответствие MTU](#)

[Показ MTU на Nexus 5000 и платформах UCS](#)

[Сквозная конфигурация](#)

[Протестируйте сквозные кадры большого размера](#)

[Буферные связанные проблемы](#)

[Проблема с драйвером](#)

[Информация об адаптере](#)

[Логический поток пакетов](#)

[Модуль ввода/вывода](#)

[Принципы проектирования](#)

[Выбор скорости порта и факторы Port Channel](#)

[Хранилище определенные проблемы](#)

[Размещение хранилища](#)

[Выбор оптимального пути](#)

[VM и совместное использование трафика гипервизора](#)

[Советы устранения неполадок](#)

[Сузьте проблему](#)

[Cisco](#)

[Встречные ограничения](#)

[Факторы уровня управления](#)

[Трафик перехвата](#)

[NetApp](#)

[VMware](#)

[Известные неполадки и усовершенствования](#)

[Обращения в центр технической поддержки](#)

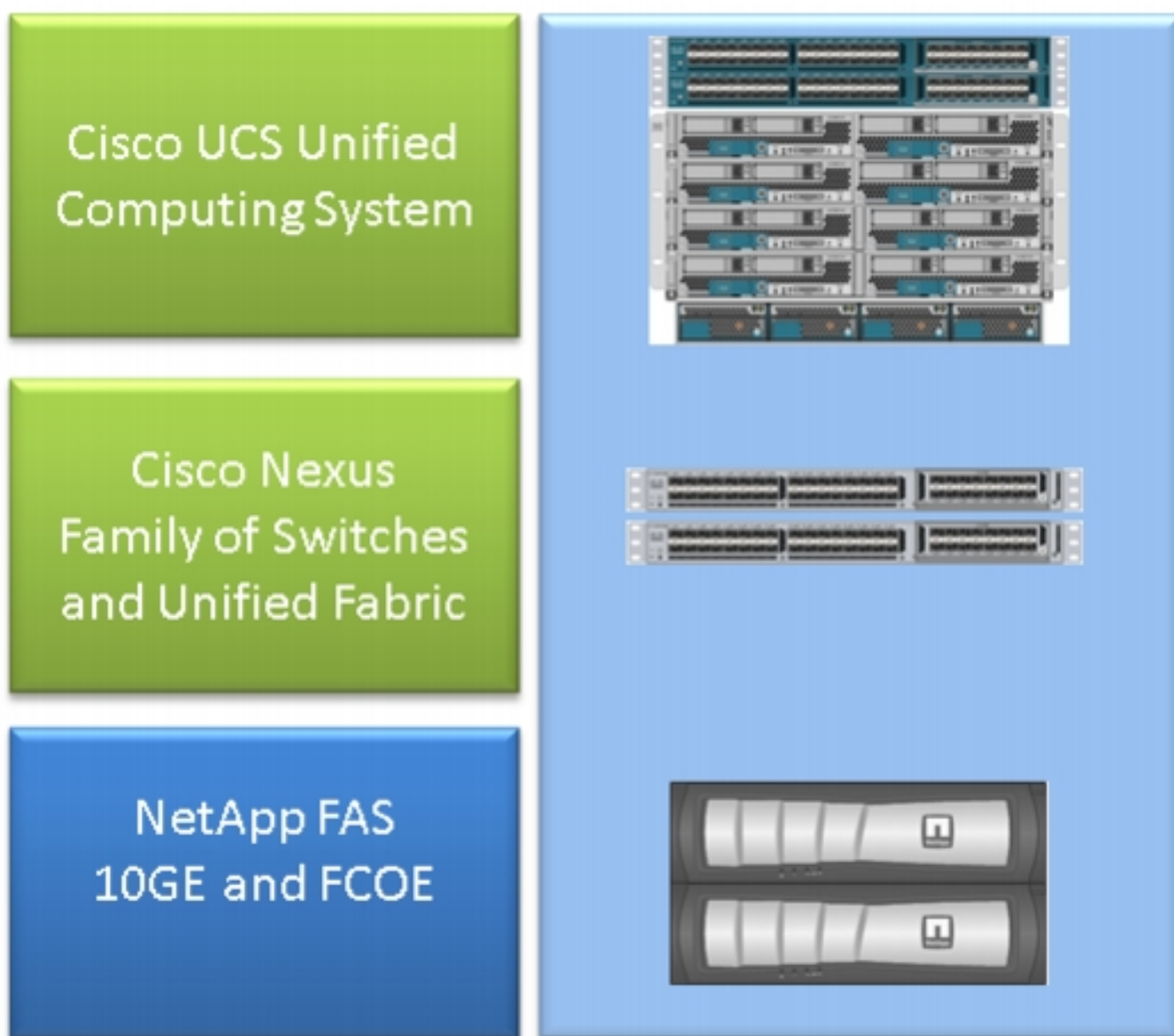
[Обратная связь](#)

Введение

Этот документ описывает общие проблемы производительности в средах FlexPod, предоставляет метод для решения проблем и предоставляет шаги смягчения. Это предназначено как отправная точка для клиентов, которые надеются устранять неполадки производительности в среде FlexPod. Этот документ был записан в результате проблем, замеченных по условию Средняя команда Центра технической поддержки (TAC) Решений в последних месяцах.

FlexPod концептуальный обзор

FlexPod состоит из компьютера системы Unified Computing System (UCS), связанного через коммутатор Nexus с системой хранения NetApp и IP - сетями.



Наиболее распространенный FlexPod состоит из UCS Cisco, который шасси серии B, связанное через Центральные устройства (FIs) к Nexus 5500, переключает на программы для работы с файлами NetApp. Другое решение, названное FlexPod Express, использует UCS шасси серии C, связанное с коммутаторами Nexus 3000. Этот документ обсуждает

наиболее распространенный FlexPod.

Вопросы производительности

В сложных средах со множественными ответственными сторонами, как, как правило, замечено в FlexPod, необходимо рассмотреть множественные аспекты для решения проблемы. Типичные проблемы производительности в Уровне 2 и IP - сетях произошли бы от:

- Пакет или потеря кадра - потеря битов данных вызывают негативное влияние на производительности приложений.
- При буферизации - если пакет или кадр проводят слишком много времени в очереди/буфере, определенные влияния производительности могли бы быть замечены приложениями, особенно в случае сетей хранения данных. Задержка, переупорядочение и проблемы нормализатора подпадают под эту категорию.
- Проблемы несоответствия MTU и фрагментация - типичная проблема, когда вы достигаете более высокой производительности. Проблемы, которые касаются фрагментации и падения несоответствия MTU этой категории.

Среда

Важно знать среду, для которой измерен уровень. Вопросы о типе хранения и протоколе, а также операционной системе (OS) влияемого сервера и местоположении, должны быть подняты для надлежащего сужения проблемы. Схема топологии, которая выделяет подключение, является обязательным минимумом.

Измерение

Необходимо знать то, что измерено и как это измерено. Определенные приложения, а также большая часть хранилища и поставщики гипервизора, предоставляют измерения некоторого вида, которые указывают на производительность/состояние системы. Эти измерения являются положительной стороной для начала в том, поскольку они не замена для большинства методик устранения проблем.

Как пример, измерение задержки хранилища Протокола NFS в гипервизоре могло бы указать, что производительность выключается, однако самостоятельно это не вовлекает сеть. В случае NFS простая проверка связи с помощью команды ping от хоста до IP - сети хранилища NFS могла бы указать, виновата ли сеть.

Срок

Эта мысль не может быть подчеркнута достаточно, особенно при открытии кейса TAC (Центра технической поддержки). Чтобы указать, что производительность является неудовлетворительной, измеренный параметр должен быть обозначен. Это включает ожидаемое и протестированное значение. Идеально, необходимо показать предыдущие данные, и методология тестирования использовала достигать тех данных.

Пример; задержка на 10 мс, достигнутая, когда протестировано, с только для записи от одиночного инициатора к Номеру единого логического блока (LUN), не могла бы быть показательной из того, что задержка, как предполагается, для полностью загружаемая система.

Проблемы производительности в FlexPod

Так как этот документ предназначен как ссылка для большинства сред FlexPod, это выделяет только большинство повторяющихся проблем, как замечено Специалистами центра технической помощи, ответственными за Решения для ЦОД.

Типичные неполадки

Проблемы, характерные для хранилища и сетей IP/Уровня 2, обсуждены в этом разделе.

Кадр и потеря пакета

Кадр и потеря пакета являются самым частым фактором, который влияет на производительность. Одно из общих мест для поиска индикаций относительно проблемы в уровне интерфейса. От Nexus 5000 или UCS Операционная система Nexus (NX-OS) CLI, введите **show interface |, сек. "подключена" | egrep ^ (Eth|fc)|discard|drop|CRC** команда. Для интерфейсов, которые подключены, это перечисляет название и сбрасывает от счетчиков и отбрасываний. Точно так же большой обзор отображен, когда вы входите, **show interface** **противостоит ошибочной** команде, которая показывает статистику ошибок для всех интерфейсов.

Мир Ethernet

Важно знать, что счетчики в не0 не могли бы указать на проблему. В определенных сценариях те счетчики, возможно, были повышены в начальной настройке или в предыдущих в рабочем состоянии изменениях. Увеличение счетчиков должно быть проверено.

Можно также собрать счетчики из уровня ASIC, который мог бы быть более показательным. В частности, для ошибки Cyclic Redundancy Checks (CRC) на интерфейсах, команда фаворита TAC для ввода является **аппаратными средствами показа внутренний carmel CRC**. Кармель является названием ASIC, ответственного за передачу уровня порта.

Подобные выходные данные могут быть взяты от FIs серии 6100, или Nexus 5600 включает для каждого порта основание. Для FI 6100, gatos ASIC, вводят эту команду:

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

Для Nexus 5600, от bigsur ASIC, вводят эту команду:

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

Команда для carmel ASIC показывает, где пакеты CRC были получены и где они были

переданы, и что еще более важно топтались ли они или нет.

И начиная с операция NX-OS Nexus 5000 и начиная с UCS является сквозной, кадры режима коммутации с неправильной Контрольной суммой фрейма (FCS) только топчутся перед передачей. Важно узнать, куда поврежденные кадры прибывают из.

```
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 |
```

Данный пример показывает топтавшие пакеты, которые прибывают из Eth 1/17 и Eth 1/18, который является каналом связи к Nexus 5000. Можно предположить, что те кадры были позже переданы вниз к Eth 1/34, такому как Eth 1/17 +, Eth 1/18 rx Топает =, Eth 1/34 tx Топает.

Подобный взгляд на Nexus 5000 показывает:

```
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 | --- | --- | --- |
```

Эти выходные данные показывают CRC, полученные на двух ссылках и отмеченные, как топают перед передачей. Для получения дополнительной информации посмотрите [руководство по поиску и устранению проблем Nexus 5000](#).

Мир Fibre Channel

Простой путь для поиска отбрасываний (discrds, ошибка, CRC, исчерпание кредита B2B) через команду **ФК счетчиков show interface**.

Эта команда, доступная на Nexus 5000 и Центральном устройстве, дает хорошую индикацию относительно того, что происходит в мире Fibre Channel.

Пример:

```
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
(...)
```

Этот интерфейс не занят, и выходные данные показывают, что не произошли никакой сброс или ошибка.

Кроме того, переходы кредита B2B от 0 были выделены; из-за идентификаторов ошибок Cisco [CSCue80063](#) и [CSCut08353](#), тем счетчикам нельзя доверять. Они хорошо работают на MDS Cisco, но не на UCS платформ Nexus5k. Также можно проверить идентификатор ошибки Cisco [CSCsz95889](#).

Так же к carmel в мире Ethernet для Fibre Channel (FC) средство ФК-Мас может использоваться. Как пример, для порта fc2/1, вводят **аппаратные средства показа внутренний ФК-Мас 2 команды statistics порта 1**. Представленные счетчики находятся в шестнадцатеричном формате.

```
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
```

Выходные данные показывают 15 сброса на вводе. С этим можно совпасть к FCP_CNTR_PIF_RX_DROP, который рассчитал к 0xf (15 в десятичном числе). Эта информация может быть снова коррелирована к FWM (Передающий Менеджеру) информация.

```
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
```

Однако это говорит администратору сумму отбрасываний и который является соответствующим номером ASIC. Получить информация о причине того отброшенного ASIC нужно делать запрос.

```
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]
```

В этом случае трафик был отброшен входным Списком контроля доступа (ACL), как правило, в мире FC - зонирование.

Несоответствие MTU

В средах FlexPod важно принять сквозное значение Maximum Transition Unit (MTU) для приложений и протоколов, где это требуется. В случае большинства сред это - Fibre Channel по Ethernet (FCoE) и кадры большого размера.

Кроме того, должна происходить фрагментация, ухудшенная производительность должна ожидаться. В случае протоколов, таких как Протокол NFS и интернет-Интерфейс scsi (iSCSI), важно протестировать и доказать сквозной Максимальный размер передаваемого блока данных (MTU) IP и Maximum Segment Size (MSS) TCP.

Устраняете ли вы неполадки кадров большого размера или FCoE, важно помнить, что обоим из них нужны однотипная конфигурация и маркировка Класса обслуживания (CoS) через среду для работы должным образом.

В случае UCS и Nexus, команда, которая полезна для проверки поинтерфейсного, на Параметр MTU группы QoS является `show queuing interface | я queuing|qos-group|MTU`.

Показ MTU на Nexus 5000 и платформах UCS

Известным аспектом и UCS и Nexus является показ MTU на интерфейсе. Эти выходные данные демонстрируют интерфейс, настроенный для организации очереди Кадров большого размера и FCoE:

```
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)
```

В то же время команда `show interface` отображает 1500 байтов:

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

Если по сравнению с `show interface` информацией о ASIC, ASIC показывает возможность MTU данного порта.

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

Это несоответствие MTU в показе ожидается на вышеупомянутых платформах и могло потенциально ввести в заблуждение новичков.

Сквозная конфигурация

Сквозная однотипная конфигурация является единственным способом гарантировать надлежащую производительность. Конфигурация кадров большого размера и шаги для стороны Cisco, а также VMware ESXi, описаны в [UCS с VMware ESXi Сквозной Jumbo Пример конфигурации MTU](#).

[Пример конфигурации UCS FCoE канала от абонента к оператору](#) показывает UCS и конфигурацию Nexus 5000. Посмотрите Приложение А в ссылочном документе для структуры основной конфигурации Nexus 5000.

[Установите Подключение FCoE для блейда UCS Cisco](#) внимание на конфигурацию UCS для

FCoE. [Nexus 5000 NPIV FCoE с FCoE NPV Подключенный Пример конфигурации UCS](#) фокусируется на конфигурации Nexus.

Протестируйте сквозные кадры большого размера

Большинство современных операционных систем предлагает способность протестировать надлежащую конфигурацию кадров большого размера с простым тестом Протокола ICMP.

Вычисление

9000 байтов - IP - заголовок без опций (20 байтов) - Заголовок ICMP (8 байтов) = 8972 байта данных

Команды в популярных операционных системах

Linux

```
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
```

Буферные связанные проблемы

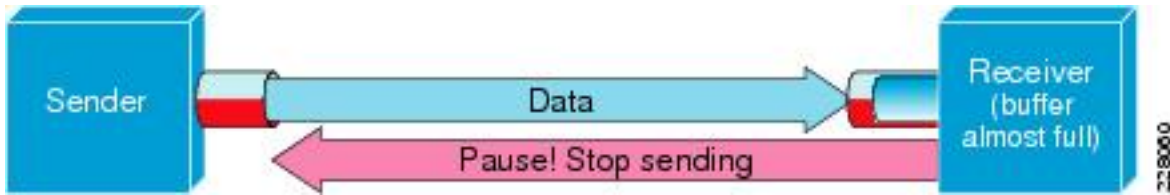
Буферизация и другие связанные проблемы задержки среди общих причин снижения производительности в среде FlexPod. Не все проблемы сообщили как основа задержки от фактических проблем буферизации, довольно много измерений могли бы указать на сквозную задержку. Например, в случае NFS, период зарегистрированного времени мог бы быть необходим к успешному чтению-записи к хранилищу и не задержке реальной сети.

Перегрузка является наиболее распространенной причиной для буферизации. В мире Уровня 2 перегрузка может вызвать буферизацию и даже хвостовые отбрасывания кадров. Во избежание отбрасываний во время периодов перегрузки были представлены фреймы паузы IEEE 802.3x и Приоритетное управление потоками (PFC). Оба полагаются на то, чтобы просить, чтобы оконечная точка держала передачи для короткого периода времени, в то время как длится перегрузка. Это может быть вызвано перегрузкой сети (сокрушите полученный с объемом данных), или потому что приоритизированный кадр должен пройти, как в случае для FCoE.

Управление потоками - 802.3x

Для проверки, который интерфейс включили управление потоками, введите команду **flowcontrol show interface**. Важно придерживаться рекомендации поставщика систем хранения в отношении включаемого управления потоками.

Рисунок, который показывает, как 802.3x работы управления потоками показан здесь.

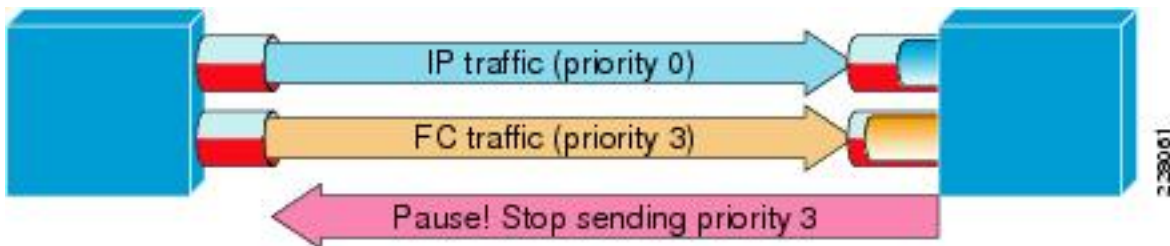


PFC - 802.1Qbb

PFC не требуется для всех настроек, но рекомендуется для большинства. Для проверки, который интерфейсам включили PFC, **приоритетное управление потоками show interface | я**, команда **On** может быть выполнена на NX-OS UCS и Nexus 5000.

Интерфейсы между FIs и Nexus 5000 должны быть видимы в том списке. В противном случае конфигурация QoS должна быть проверена. QoS должно быть последовательно сквозной для использования преимуществ PFC. Для проверки, почему PFC не подходит на определенном интерфейсе, введите **show system внутренний интерфейс "Ethernet" журнала dcbx x/y** команда для получения ЦОД, Соединяющего Протокол обмена Возможностей (DCBX) журнал.

Рисунок, который показывает, как фреймы паузы работают с PFC, показывают здесь.



Команда **приоритетного управления потоками show interface** позволяет администратору наблюдать поведение класса на QoS приоритетных фреймов паузы.

Пример:

```
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)
```

Эти выходные данные показывают, что во втором классе устройство просто передавало (TX) кадр PPP.

В этом случае Ethernet 1/1 является портом, стоящим перед IOM и в то время как полному порту не включают PFC, это могло бы обработать кадры PPP для портов FEX.

```
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
```

В этом случае интерфейсы FEX включены.

```
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
```

Порты FEX, которые включены, могут быть также проверены через показ **fex X** **подробностей** , где X номер стойки.

```
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
```

См. эти документы для получения дополнительной информации о механизмах паузы.

- [Fibre Channel по операциям Ethernet](#)
- [Унифицированный оптоволоконный Fibre Channel Описания технологических решений по Ethernet \(FCoE\)](#)

Организация очереди сброса

И Nexus 5000 и NX-OS UCS отслеживают входной сброс из-за организации очереди на на основании группы QOS. Пример:

```
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)
```

Входной сброс *должен* произойти только в очередях, которые настроены для разрешения отбрасываний.

Входной сброс организации очереди может произойти из-за этих причин:

- Коммутируемый анализатор для портов (SPAN) / Контролирующий сеанс включил на некоторых интерфейсах (см. идентификатор ошибки Cisco [CSCur25521](#)),

- Обратное давление от другого интерфейса, фреймы паузы, как правило, замечаются, когда включено
- Трафик плыл на плоскодонке к ЦП

Проблема с драйвером

Cisco предоставляет два драйвера операционной системы для UCS, enic и fnic. Enic ответственен за Подключение по технологии Ethernet, и fnic ответственен за подключение Fibre Channel и FCoE. **Очень важно**, чтобы enic и fnic драйверы были точно как заданы в [матрице совместимости UCS](#). Проблемы, представленные неправильными драйверами, колеблются от потери пакета и добавленной задержки к более длинному процессу загрузки или полному отсутствию подключения.

Информация об адаптере

Предоставленный Cisco адаптер может предоставить хорошее измерение о трафике, который передают, а также отбрасывания. Данный пример показывает, как соединиться с шасси X, сервером Y и адаптером Z.

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

Отсюда, администратор может войти к Центру мониторинга для Производительности (MCP) в средство.

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

Средство MCP позволяет вам контролировать использование трафика на логический интерфейс (LIF).

```
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
-----
```

Шасси 1, разъедините 1, и адаптер 1 имеет две Карты Виртуального сетевого интерфейса (VNICs), привязанный к виртуальным интерфейсам (Действительная Ethernet или Действительный Fibre Channel) 834 и 836. У тех есть номера 2 и 3. Статистика для LIF 2 и 3 может быть проверена как показано здесь:

```
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

```

Следует отметить, что администратору UCS предоставляют общее количество и дельту (между двумя последующим выполнением lifstats) столбцы, а также текущий трафик на LIF и информация о любых ошибках, которые, возможно, произошли.

Предыдущий пример показывает интерфейсы без любых ошибок с очень маленькой загрузкой. Данный пример показывает другой сервер.

```

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

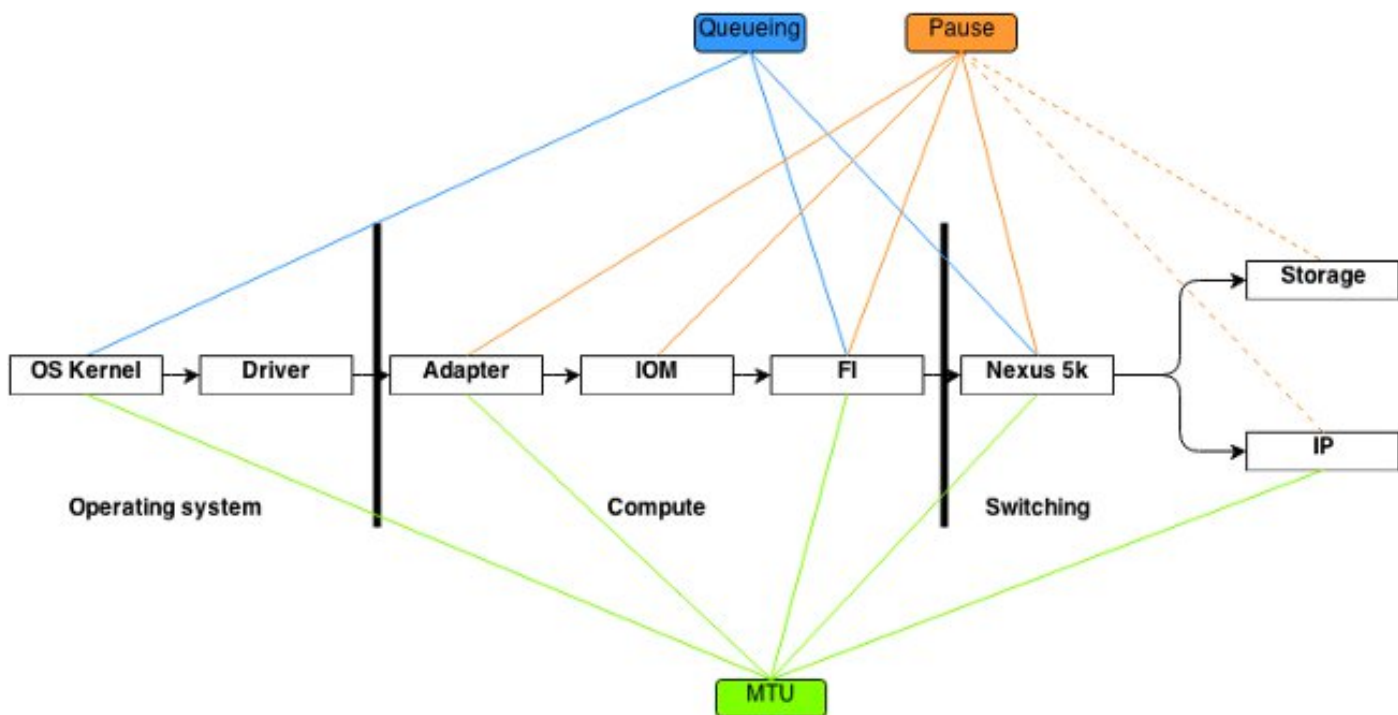
784.185kbps    Rx rate

```

Два содержательных бита информации показывают, что 96 кадров были отброшены адаптером из-за отсутствия буфера или буферизации отключенного и дополнительно обрабатываемые сегменты Разгрузки сегмента TCP (TSO).

Логический поток пакетов

Схема, показанная здесь, выделяет логический поток пакетов в среде FlexPod.



Эта схема предназначена как отказ компонентов кадр, через который проходят на пути через среду FlexPod. Это не отражает сложность ни одного из блоков и является просто способом запомнить, где определенные характеристики должны быть настроены и проверены.

Модуль ввода/вывода

Как показано в логической схеме потока пакетов, Модуль ввода/вывода (IOM) является компонентом посреди всей связи, которая проходит UCS. Для соединения с IOM в шасси X, введите **подключение iom x** команда.

Вот несколько других полезных команд:

- Информация о топологии - **программное обеспечение show platform [woodside|redwood] команда sts** показывает топологическую информацию с точки зрения IOM.


```

# show platform software worldwide loss

```

Port	SND			Port Extra Drop	S8 Loss Counters	Cos_u														
	Tx Pack	Rx Pack	Errors			Counters	COS													
0-NI2	0	82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0-NI23	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Из-за пути базовая инфраструктура работает, счетчики показывают только для интерфейсов, которые испытали любую потерю промежуточное выполнение этих двух команд. В данном примере вы видите, что интерфейс NI2 получил 82 фрейма паузы и что 28 фреймов паузы были переданы для взаимодействия через интерфейс NI23, который вы знаете, присоединен к блейду 3.

Принципы проектирования

FlexPod обеспечивает гибкую конфигурацию и устанавливать хранилища и сети передачи данных. С гибкостью также прибывает дополнительные проблемы. Жизненно важно придерживаться документов оптимальных методов и Cisco проверила дизайн (CVD):

- CVD - [Руководство по развертыванию FlexPod](#)
- Оптимальные методы системы хранения NetApp (не определенный для Flexpod) - [система Cisco UCS \(UCS\) Опции Подключения ресурсов хранения и Оптимальные методы с системой хранения NetApp](#)

Выбор скорости порта и факторы Port Channel

Типичная проблема, замеченная инженерами TAC, является избыточным использованием ссылок из-за выбора 1 Ethernet Gbit вместо 10 Ethernet Gbit, на которые ссылаются в документах оптимального метода. Как резкий пример, производительность **единого потока** не будет лучше на десять 1 ссылка Gbit по сравнению с одними 10 ссылками Gbit. В канале порта единый поток может пробежаться через одно соединение.

Для обнаружения, какое распределение нагрузки метода используется на Nexus и/или NX-OS Fi, введите команду **show port-channel load-balance**. Администратор может также узнать, какой интерфейс в канале порта будет выбран в качестве исходящего интерфейса для пакета или кадра. Простой пример кадра на VLAN49 между двумя хостами показывают здесь:

```

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
src-mac: 70ca.9bce.ee24

```

Хранилище определенные проблемы

Проблемы, обсужденные ранее, характерны и для данных и для сетей хранения данных. Ради полноты также упомянуты проблемы производительности, определенные для Сети хранения данных (SAN). Протоколы хранения были созданы с упругостью, и mutli-соединение-каналом все еще увеличены. С появлением технологий, таких как Асимметричное присвоение логического модуля (ALUA) и Многопутевой IO (MPIO), большая гибкость и опции представлены администраторам.

Размещение хранилища

Другое рассмотрение является размещением хранилища. Дизайн FlexPod диктует то хранилище, должен быть подключен на коммутаторах Nexus. Непосредственно подключенная система хранения не соответствует CVD. Если оптимальные методы применены, дизайны с непосредственно подключенной системой хранения поддерживаются. В то же время теми дизайнами не является строго FlexPod.

Выбор оптимального пути

Это - технически не проблема Cisco, поскольку большинство тех опций очевидно для устройств Cisco. Это - типичная проблема, чтобы выбрать и придерживаться оптимального пути. Современному Устройству определенному модулю (DSM) можно предоставить разнообразные пути и потребности выбрать оптимальную (s), на основе определенных критериев для обеспечения упругости и распределения нагрузки. Этот снимок экрана показывает четыре пути, доступные DSM NetApp для опций распределения нагрузки и Microsoft Windows.

The screenshot displays a storage management interface with a table of paths and a dialog box for DSM properties.

Disk ID	Path ID	Operational State	Admin State	Initiator Name	Initiator Address
Disk0	01000101	Active/Optimized	Enabled	com.ciscosystem...	20:00:00:25:b5:00:a...
Disk0	02000002	Active/Non-Optimized	Enabled	com.ciscosystem...	20:00:00:25:b5:00:b...
Disk0	01000001	Active/Optimized	Enabled	com.ciscosystem...	20:00:00:25:b5:00:a...
Disk0	02000102	Active/Non-Optimized	Enabled	com.ciscosystem...	20:00:00:25:b5:00:b...

The dialog box, titled "Data ONTAP(R) DSM Properties", has tabs for "Data ONTAP DSM", "MPIO", and "License Information". Under the "MPIO" tab, the "Default Load Balance Property" section contains the following options:

- Auto Assign
- Failover Only
- Round Robin
- Round Robin with Subset
- Least Weighted Paths
- Least Queue Depth

Рекомендуемые настройки должны быть выбраны на основе обсуждения с поставщиком систем хранения. Те параметры настройки могли бы влиять на проблемы производительности. Типичный тест, который ТАС мог бы попросить, чтобы вы выполнили,

является тестом чтения-записи через только матрицу А или матрицу В. Это, как правило, позволяет вам сужать проблемы производительности к ситуациям, обсужденным в разделе "Типичных проблем" этого документа.

VM и совместное использование трафика гипервизора

Эта точка является определенной для вычислить компонента, независимо от поставщика. Простой способ для устанавливания сети хранения данных для гипервизоров с вычислить точки зрения должен создать два Host Bus Adapter (HBA), один для каждого Волокна, и выполнить и загрузочный трафик LUN и трафик СХД Виртуальной машины (VM) по тем двум интерфейсам. Всегда рекомендуется разделить загрузочный трафик LUN и трафик СХД VM. Это обеспечивает лучшую производительность и дополнительно позволяет логическое разделение между двумя видами трафика. Посмотрите раздел "Известных неполадок" для примера.

Советы устранения неполадок

Сузьте проблему

Как в случае любого быстрого устранения проблем, очень важно сузить проблему и задать правильные вопросы.

- Какие устройства/приложения/VM (/не) влияются?
- Какой контроллер хранения (/не) влияется?
- Какие пути (/не) влияются?
- Как часто делает проблему (/не), появляются?

Cisco

Встречные ограничения

В интерфейсе этого документа обсуждены счетчики организации очереди ASIC. Счетчики также высказывают мнение в момент времени, таким образом, важно контролировать увеличение счетчиков. Определенные счетчики не могут быть очищены дизайном. Например, carmel ASIC, упомянутый ранее.

Для предоставления резкого примера присутствие CRC или сброса на интерфейсе не могло бы быть идеальным, но можно было бы ожидать, что их значения являются ненулевыми. Счетчики, возможно, повысились в некоторый момент вовремя, возможно во время перехода или начальной настройки. Следовательно важно обратить внимание на увеличение счетчиков и когда был прошлый раз, они были очищены.

Факторы уровня управления

В то время как полезно рассмотреть счетчики, важно знать, что определенные проблемы

плоскости данных не могли бы найти легкое отражение к счетчикам уровня управления и программным средствам. Как резкий пример, ethalyzer очень полезный инструмент, который доступен и на UCS и на Nexus 5000. Однако это может только перехватить трафик уровня управления. Перехват трафика - то, что TAC часто запрашивает, особенно когда не ясно, где лежит вина.

Трафик перехвата

Надежный трафик перехватывает взятый, конечные хосты могут пролить свет на проблему производительности и сузить его довольно быстро. И Nexus 5000 и UCS предлагают SPAN трафика. В частности опции UCS ОХВАТЫВАНИЯ определенных HBA и оптоволоконных сторон полезны. Для узнавания больше о возможностях перехвата трафика при мониторинге сеанса на UCS посмотрите эти ссылки:

- [Анализ трафика UCS для физического и виртуальных адаптеров](#) (видео)
- [Руководство конфигурации GUI Cisco UCS Manager - контролирующей трафик](#)

NetApp

NetApp предлагает полный набор утилит для устранения проблем их контроллеров хранения, среди них:

- perfstat - очень полезная утилита, как правило, работайте для персонала службы технической поддержки NetApp
- systat - предоставляет сведения о том, насколько занятый программа для работы с файлами и что программа для работы с файлами делает - [Вспомогательная библиотека NetApp](#)

Существуют среди наиболее распространенных команд:

- `sysstat -x 2`
- `sysstat -M 2`

Вот некоторые вещи искать в выходных данных `sysstat-x 2`, которые могли бы указать на перегруженный массив NetApp или диски:

- Длительный **CP ty** столбец с большим количеством из: или **F**
- Длительный **HDD util** столбец выше **20%**

Эта статья описывает, как настроить NetApp: [Оптимальные методы Хранилища Ethernet NetApp](#) .

- Маркирование VLAN
- VLAN Trunking
- Jumbo MTU
- Хеширование IP
- Отключение управления потоком

VMware

ESXi предоставляет доступ Secure Shell (SSH), через который можно устранить неполадки. Среди большинства полезных инструментов, предоставленных администраторам, esxtop и

perfmon.

- esxtop - во многом как вершина Linux/BSD это позволяет пользователям контролировать связанные параметры производительности в реальном времени
[Использование esxtop для определения проблем производительности хранилища для ESX / ESXi](#)
- perfmon - позволяет пользователям устранять неполадки Виртуальных машин (VM) Microsoft Windows
[Сбор данных журнала Windows Perfmon для диагностирования проблем производительности виртуальной машины](#)
- Соберите диагностическую связку (bundle) на ESXi - [Собирающая диагностическая информация для VMware ESX/ESXi использование vSphere Клиента \(653\)](#)
- Требование Распределения нагрузки VMware vSwitch для Cisco серверы Серии B - [Маршрут на основе хэша IP не поддерживается с блейд-серверами Cisco UCS B200 M1/M2, которые используют UCS Центральные устройства серии 6100](#)

Известные неполадки и усовершенствования

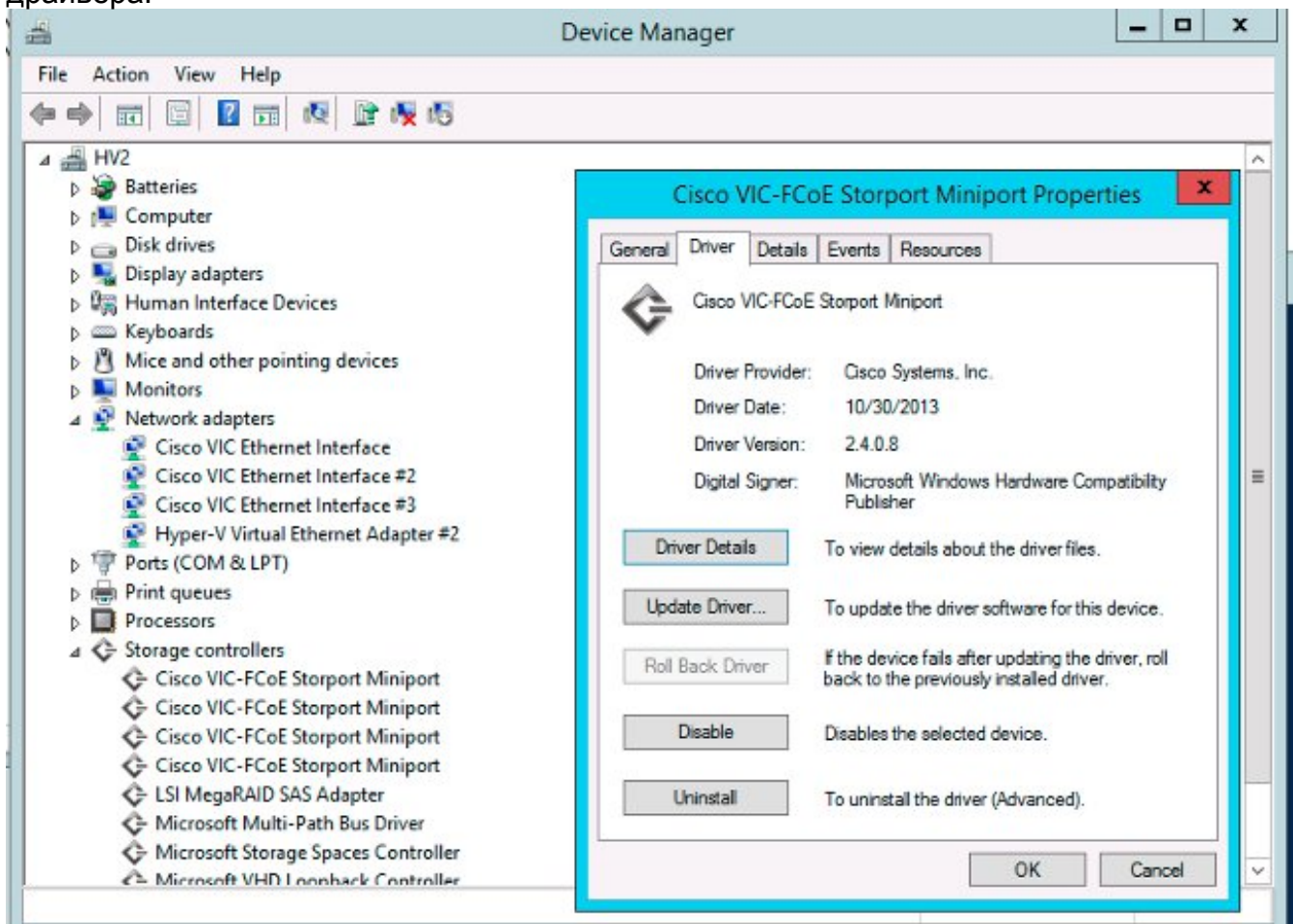
- Идентификатор ошибки Cisco [CSCuj86736](#) - с пассивными twinax ошибками CRC кабелей может увеличиться. Когда Nexus 5000 не оптимизирует DFE, это вызвано. Введите аппаратные средства показа внутренняя `show csmel` команда, чтобы проверить, что "Глазной параметр" высоты выше 100 милливольт. Это было исправлено в Версиях 5.2 (1) N1 (7) и 7.0 (4) N1 (1).
- Идентификатор ошибки Cisco [CSCuo76425](#) - подобный предыдущему дефекту и также существует на центральных устройствах UCS. Это исправлено в Выпуске 2.2 (3a).
- Идентификатор ошибки Cisco [CSCuo76425](#) - то же как дефект [CSCuj86736](#) за исключением Центрального устройства UCS.
- Идентификатор ошибки Cisco [CSCup40056](#) - ошибка синхронизации, вызванная путем совместного использования загрузочного трафика с трафиком VM, описанным в [Системной Виртуальной машине Унифицированных вычислений Оперативных Сбоев Миграции с Действительными Адаптерами Fibre Channel](#).
- На медленное обнаружение дренажа и предотвращение - очень часто FC и FCoE влияет медленный дренаж. Выпуск 7.0 (0) N1 (1) NX-OS представляет средства обнаружить и избежать его. Узнайте больше о функции в [Руководстве по конфигурации Интерфейсов NX-OS серии 5500 Cisco Nexus](#) и [Медленном Обнаружении устройств Дренажа и Предотвращении перегрузки](#).
- Идентификатор ошибки Cisco [CSCuj81245](#) - ограничение существует в основанных картах PALO (VIC1240 и другие), который вызывает прерывания FC.
- Идентификатор ошибки Cisco [CSCuh61202](#) - после обновления к Выпуску 2.1 (3), прерываниям FC микропрограммного обеспечения UCS и множественным другим проблемам может быть замечен.
- Идентификатор ошибки Cisco [CSCtw91018](#) - соединение Параметров MTU для VNICs на одиночном, основанном на PALO адаптере может вызвать исчерпание ресурсов для некоторых классов трафика.
- Идентификатор ошибки Cisco [CSCuq40256](#) - заставит PFC быть отключенным на ссылках от Центрального устройства вниз к адаптерам сервера. Это вызовет множество проблем, которые запускаются с прерываний Fibre Channel, и Неисправные кадры

сообщили относительно стороны хранилища. О разъединениях хранилища и других проблемах производительности можно было бы сообщить.

Обращения в центр технической поддержки

Во многих случаях инженер ТАС попросит, чтобы вы собрали некоторую основную информацию, прежде чем сможет быть запущено расследование.

- Схема топологии - который включает номера портов и скорости линии, абсолютно необходимые.
- Техническая поддержка UCSM - [Визуальное Руководство для сбора файлов Технической поддержки \(В и серии С\)](#).
- Техническая поддержка шасси UCS для одного шасси, которое испытывает проблемы - видит предыдущую ссылку.
- И техническая поддержка Nexus 5000 и любые другие сетевые устройства между UCS и [выходными данными NetApp - Redirecting техподдержки показа детализируют команду](#).
- Выходные данные команды **show queueing interface** на обоих из FIs. `sysstat -m 2`
- Версии драйвера узла на ESXi выполняют - вводят эти команды: `vmkload_mod-s enicvmkload_mod-s fnic`
- Linux - `sysstat -m 2`
- Windows - проверяют версию драйвера в "менеджере устройств". Пример из Окна 2012 R2 показывает три Интерфейса Ethernet VIC Cisco и четыре интерфейса VIC FCoE miniport (ответственный также за Fibre Channel, не только FCoE) и Выпуск 2.4.0.8 fnic драйвера.



Обратная связь

Используйте кнопку feedback для обеспечения отзыва об этом документе или опыте. Мы будем постоянно обновлять этот документ, поскольку разработки происходят и после того, как получен отзыв.