

FlexPod普通的性能问题

Contents

[Introduction](#)

[FlexPod概念性概述](#)

[性能注意事项](#)

[环境](#)

[测量](#)

[基准](#)

[在FlexPod的性能问题](#)

[常见问题](#)

[帧和信息包丢失](#)

[MTU 不匹配](#)

[在连结5000和UCS平台的MTU显示](#)

[端到端配置](#)

[测试端到端巨型帧](#)

[缓冲相关问题](#)

[驱动器问题](#)

[适配器信息](#)

[逻辑信息包流](#)

[输入-输出模块](#)

[设计注意事项](#)

[端口速度选择和端口通道考虑](#)

[存贮特定问题](#)

[存贮安置](#)

[最佳路径选择](#)

[VM和Hypervisor流量共享](#)

[故障排除提示](#)

[缩小问题](#)

[Cisco](#)

[计数器限制](#)

[控制层面考虑](#)

[捕获数据流](#)

[NetApp](#)

[VMware](#)

[已知问题和增进](#)

[TAC案例](#)

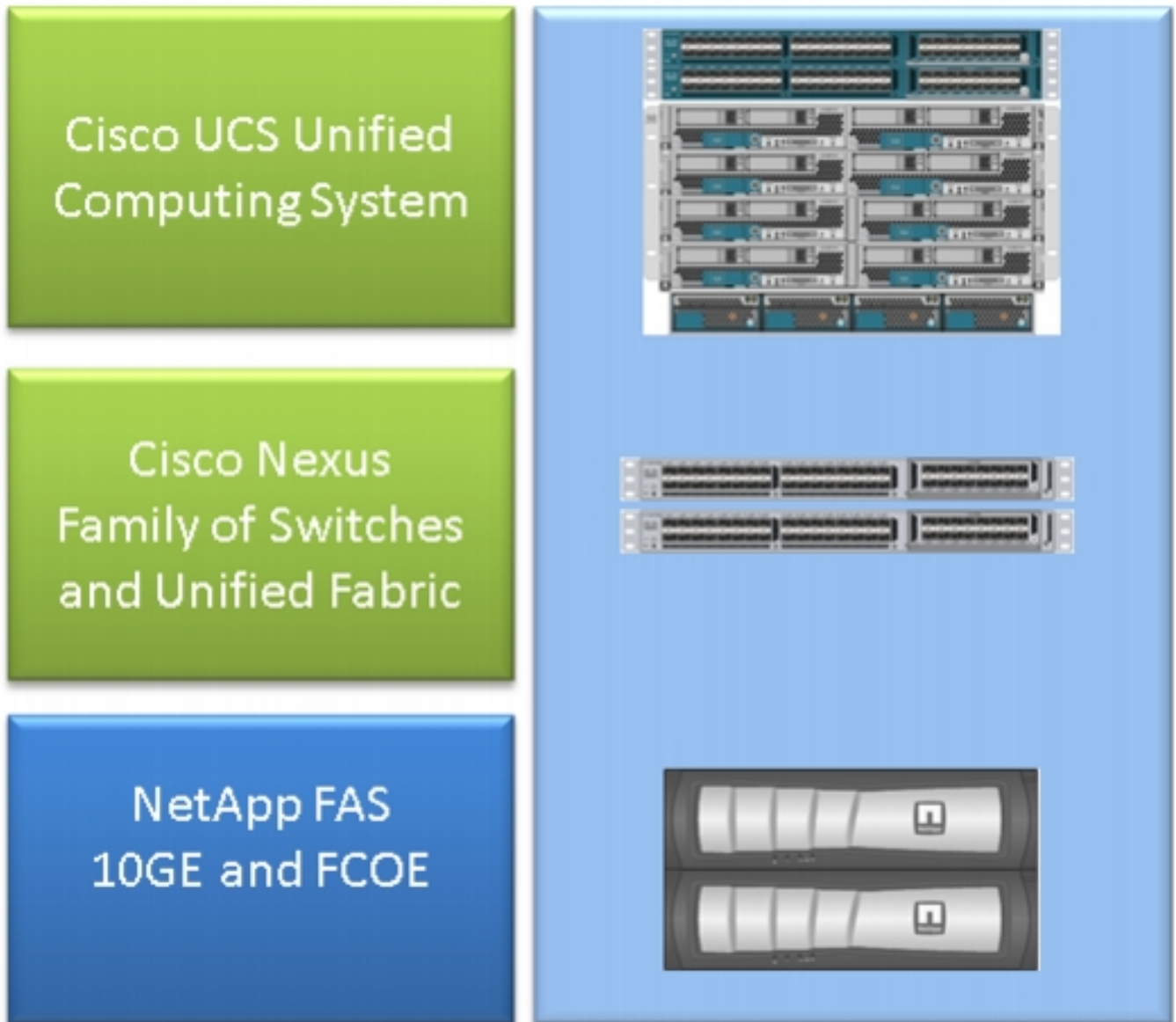
[Feedback](#)

Introduction

本文在FlexPod环境描述普通的性能问题，提供方法排除问题故障，并且提供缓解步骤。打算作为在FlexPod环境里查找排除性能故障的用户的起始点。本文被写作了由于数据中心解决方案技术支持中心(TAC)小组近几个月来看到的问题。

FlexPod概念性概述

FlexPod包括通过连结交换机被连接的统一计算系统(UCS)计算机到NetApp存贮和IP网络。



最普通的FlexPod包括通过结构被连接的Cisco UCS B系列机箱互联(FIs)对连结5500交换机到NetApp锉刀。另一个解决方案，称为FlexPod Express，使用一个UCS C系列机箱被连接到连结3000交换机。本文讨论最普通的FlexPod。

性能注意事项

在与多个负责的当事人的复杂环境如典型地被看到在FlexPod，您需要考虑多个方面为了排除问题故障。在第2层的典型的性能问题和IP网络将源于：

- 信息包或帧丢失-位损失数据引起对应用程序性能的负面作用。
- 缓冲-，如果信息包或帧特别是在存储网络的情况下花费在某些性能指示也许看到由应用程序的队列/缓冲区的许多时刻。潜伏期，重新命令和规整器问题属于此类别。
- MTU不匹配问题和分段-常见问题，当您到达高性能。与分段和MTU不一致关连的问题在此类别符合。

环境

认识性能被测量的环境是重要的。应该上升关于存贮类型的问题和协议，以及受影响的服务器的操作系统(OS)和位置，适当地缩小问题。概述连接的拓扑图是最小值。

测量

您需要知道什么被测量，并且它如何被测量。某些应用程序，以及多数存贮和hypervisor供应商，提供指示系统的性能/健康某个排序的评定。因为他们不是多数故障排除方法的一个替代品这些评定是开始的一个好观点在。

为例，在hypervisor的一网络文件系统存贮潜伏期测量也许表明性能断开，然而独自地不牵连网络。一旦NFS，从主机的简单PING到NFS存贮IP网络也许指示网络是否是责备。

基准

特别是当您开TAC案例时，此点不可能强调足够。为了表明性能是令人不满的，被测量的参数需要指示。这包括期望的和测试的值。理论上讲，您应该显示早先用于的数据和测试方法达到该数据。

为例;10ms达到的潜伏期，当测试，与一只读从单个发起者到逻辑部件号(LUN)，也许不是预示的什么潜伏期应该是为一个充分地加载的系统。

在FlexPod的性能问题

因为本文作为多数的参考打算FlexPod环境，略述仅常见的问题如看到由TAC小组负责对数据中心解决方案。

常见问题

问题普通对存贮和IP/Layer 2网络在此部分讨论。

帧和信息包丢失

帧和信息包丢失是最常见的要素该影响性能。寻找问题的征兆的其中一个普通的地方在界面水平。从连结5000或UCS连结操作系统的(NX-OS) CLI，请输入**show interface|秒“启用”|egrep ^**

(Eth|fc)|丢弃|丢弃|crc命令。对于是UP的接口，它列出名字并且丢弃计数器和丢包。同样地，巨大概述显示，当显示所有接口的错误统计数据时您输入error命令时show interface的计数器。

以太网世界

知道是重要的在non-0的计数器也许不指示问题。在某些情况下那些计数器也许已经被上升了在初始建立或在早先可操作的更改。应该监控计数器的增量。

一能也采集从ASIC级别的计数器，也许是更加预示的。特别地，对于在接口的循环冗余校验(CRC)错误，TAC偏爱的命令进入是show hardware内部carmel crc。Carmel是ASIC的名字负责对端口级转发。

相似的输出可以从6100系列FIs或连结被采取5600交换机在每个端口。对于FI 6100，gatos ASIC，输入此命令：

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

对于从bigsur ASIC的连结5600，请输入此命令：

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

carmel的ASIC命令显示CRC信息包哪里收到了，并且他们转发了到的地方，并且更加重要地是否他们重踏。

因为连结5000和UCS NX-OS操作是切入直通，与不正确帧校验Sequence(FCS)的交换模式帧在转发前只重踏。发现是重要的损坏的帧何处来自。

```
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 |
```

此示例显示来自Eth 1/17和Eth 1/18，是上行链路对连结5000的重踏的信息包。一个人能假设，那些帧稍后被发送了下來到Eth 1/34，例如Eth 1/17 + 1/18 rx重踏= Eth 1/34 tx重踏的Eth。

在连结5000的一相似的查看显示：

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

此输出显示CRC接收在两条链路和被标记作为重踏在转发前。欲知更多信息，请参阅[连结5000故障排除指南](#)。

光纤信道世界

简单方法寻找丢包(discrds、错误、CRC，B2B信用值耗尽)是通过show interface计数器fc命令。

此命令，可用在连结5000和结构互连，给予什么的一个好征兆在光纤信道世界发生。

例如：

```
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
(...)
```

此接口不是繁忙的，并且输出表示，丢弃或错误没有发生。

另外，B2B从0的信用值转变突出显示了;由于Cisco Bug ID [CSCue80063](#)和[CSCut08353](#)，那些计数器不可能委托。他们良好工作在Cisco MDS，但是不在Nexus5k平台UCS。并且您能验证Cisco Bug ID [CSCsz95889](#)。

同样于在以太网世界的carmel光纤信道(FC)的可以使用FC MAC设备。例如，对于端口fc2/1，请输入statistics命令show hardware内部FC MAC 2的端口1。被提交的计数器以十六进制格式。

```
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

输出显示在输入的15丢弃。这可以被匹配到计数对0xf的FCP_CNTR_PIF_RX_DROP (15在十进制)。此信息可以与FWM (转发管理器)信息再关联。

```
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
```

然而，此tells管理员，并且是对应的ASIC编号的相当数量丢包。关于那的原因的获得信息下降了ASIC需要被查询。

```
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]
```

在这种情况下，数据流由入口访问控制表(ACL)降低，典型地在FC世界-区域。

MTU 不匹配

在FlexPod环境里适应需要的应用程序和协议的端到端最大转换单元(MTU)是重要的设置。一旦多数环境，这是在以太网(FCoE)和巨型帧的光纤信道。

另外，应该分段发生，下降的性能将预计。在协议的情况下例如网络文件系统和Internet Small Computer System Interface (iSCSI)，测试和证明端到端IP最大传输单元(MTU)和TCP最大分段尺寸(MSS)是重要的。

您是否排除巨型帧或FCoE故障，请记住两个那些需要指示在环境间的一致配置和业务类别(CoS)为了正常运行。

一旦UCS和连结，是有用验证单个接口的一个命令，每QOS组MTU设置是显示排队接口|我排队|QOS组|MTU。

在连结5000和UCS平台的MTU显示

UCS和连结的已知方面是MTU显示在接口的。此输出展示被配置的一个接口排队巨型帧和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)
```

同时，**show interface**命令显示1500个字节：

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

如果与carmel ASIC信息比较，ASIC显示一个特定端口的MTU功能。

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

在显示的此MTU不匹配在上述平台预计，并且可能潜在误导初学者。

端到端配置

端到端一致的配置是保证适当的性能的唯一方法。Cisco边的巨型帧配置和步骤，以及VMware ESXi，在[与VMware ESXi端到端庞然大物MTU配置示例的UCS描述](#)。

[UCS FCoE上行链路配置示例](#)显示一种UCS和连结5000配置。请参阅在参考文档的附录A关于一种基本的连结5000配置的分级显示。

[设置Cisco UCS前端重点的FCoE连接](#)在FCoE的UCS配置。[连结与FCoE NPV的5000个NPIV FCoE附有了UCS在连结配置的配置示例](#)重点。

测试端到端巨型帧

多数现代日操作系统提供能力测试与一个简单的互联网控制消息协议(ICMP)测试的一种适当的巨型帧配置。

计算

没有选项(20个字节)的9000位元组IP表头- ICMP报头(8个字节) = 8972字节的数据

In命令常见的操作系统

Linux

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

微软视窗

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

ESXi

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

缓冲区相关问题

缓冲的和其他潜伏期相关问题是普通的性能降低原因中在FlexPod环境里。不是作为潜伏期被报告的的所有的问题源于实际缓冲问题，相当多评定也许指示端到端潜伏期。例如，一旦NFS，报告的时间时间期也许顺利地是需要的读/写到存贮和不实际网络传输时延。

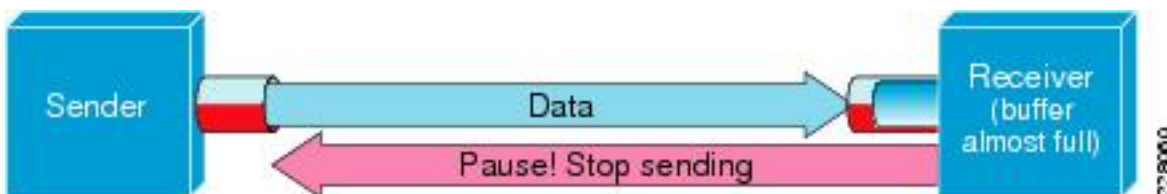
拥塞是缓冲的多数常见原因。在第2层世界，拥塞能导致缓冲和帧甚而尾部丢弃。在拥塞期间，为了避免丢包，IEEE 802.3x引入暂停帧和优先级流控制(PFC)。当拥塞持续时，两个依靠请求终点一段

时间里拿着发射。这可以由网络拥塞造成(请淹没接收的与相当数量数据)或，因为一个优先安排的帧需要通过，正如在FCoE的论点。

流控制- 802.3x

为了验证哪些接口有被启用的流控制，请输入flowcontrol命令的show interface。关于是的流控制遵从存储供应商的推荐启用的是重要的。

显示的例证802.3x如何流控制工作显示得这里。

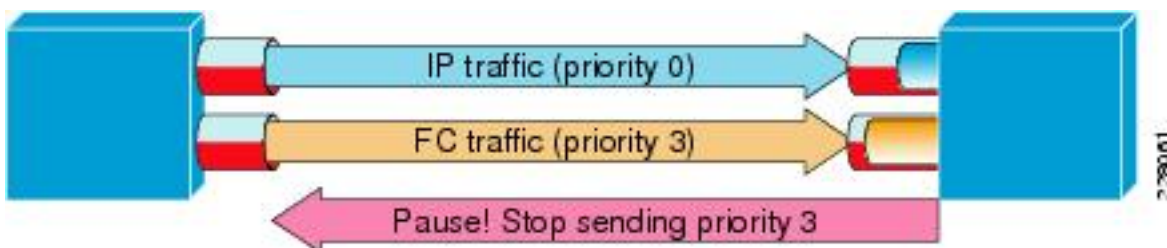


PFC -802.1Qbb

PFC没有对于所有设置是必需的，然而为多数是推荐的。为了验证哪些接口有被启用的PFC，show interface优先级流控制命令在UCS的NX-OS和连结5000可以运行。

FIs和连结5000之间的接口应该是可视的在该列表。否则，QoS配置需要被验证。QoS需要一致端到端为了利用一等兵为了检查PFC为什么在一个特殊接口不出现，输入show system内部dcbx日志以太网接口x/y命令为了获得数据中心桥接功能开关协议(DCBX)日志。

显示的例证暂停帧如何与PFC一起使用显示得这里。



show interface优先级流控制命令允许管理员观察优先级暂停帧每QoS类行为。

示例如下：

```
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)
```

此输出表示，在第二个类，设备传输(TX) PPP帧。

在这种情况下，以太网1/1是面对国际移民组织的端口，并且，而整体端口不会有被启用的PFC，也许处理FEX端口的PPP帧。

```
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
```

在这种情况下，FEX接口是包含的。

```
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
```

是包含的FEX端口可以也被检查通过显示fex x详细资料X哪里是底盘数字。

```
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
```

请参阅这些文件关于暂停机制的更多信息。

- [在以太网操作的光纤信道](#)
- [在以太网\(FCoE\)的统一的结构空白白皮书光纤信道](#)

排队的丢弃

连结5000和UCS NX-OS记录入口丢弃由于排队在a每个QOS组基本类型。例如：

```
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)
```

入口丢弃在配置允许丢包的队列应该仅发生。

入口排队的丢弃能发生由于这些原因：

- 在某些启用的交换端口分析器(SPAN) /Monitoring会话接口(请参阅Cisco Bug ID [CSCur25521](#))
- 从另一个接口的后向压力，暂停帧典型地被看到，当启用
- 数据流被踢对CPU

驱动器问题

Cisco为UCS提供两操作系统的驱动程序，enic和fnic。Enic对以太网连接负责，并且fnic对光纤信道和FCoE连接负责。重要的是非常enic和fnic驱动程序正确地是在[UCS互操作性表上指定](#)。不正确驱动程序引入的问题范围自信息包丢失和被添加的潜伏期到更长的启动流程或完成缺乏连接。

适配器信息

一台Cisco提供的适配器能提供关于通过的数据流的一好测量，以及丢弃。此示例显示如何连接到机箱x，服务器Y和适配器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.
```

从这里，管理员能登陆到性能(MCP)设备监控中心。

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

MCP设备允许您监控数据流使用方法每个逻辑接口(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

机箱1，切断1，并且适配器1有两个虚拟网络网络界面卡(VNICs)连结与虚拟接口(虚拟以太网或虚拟光纤信道) 834和836。那些有第2和3。 LIF的2和3统计数据可以被检查如显示这里：

```
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

```

请注意UCS的管理员带有总数和Delta (在lifstats之间的两随后的执行)列以及当前数据流负载每LIF和信息关于也许已经生成的所有错误。

前一个示例显示接口不出任何错误与非常小的负荷。此示例显示一个不同的服务器。

```

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

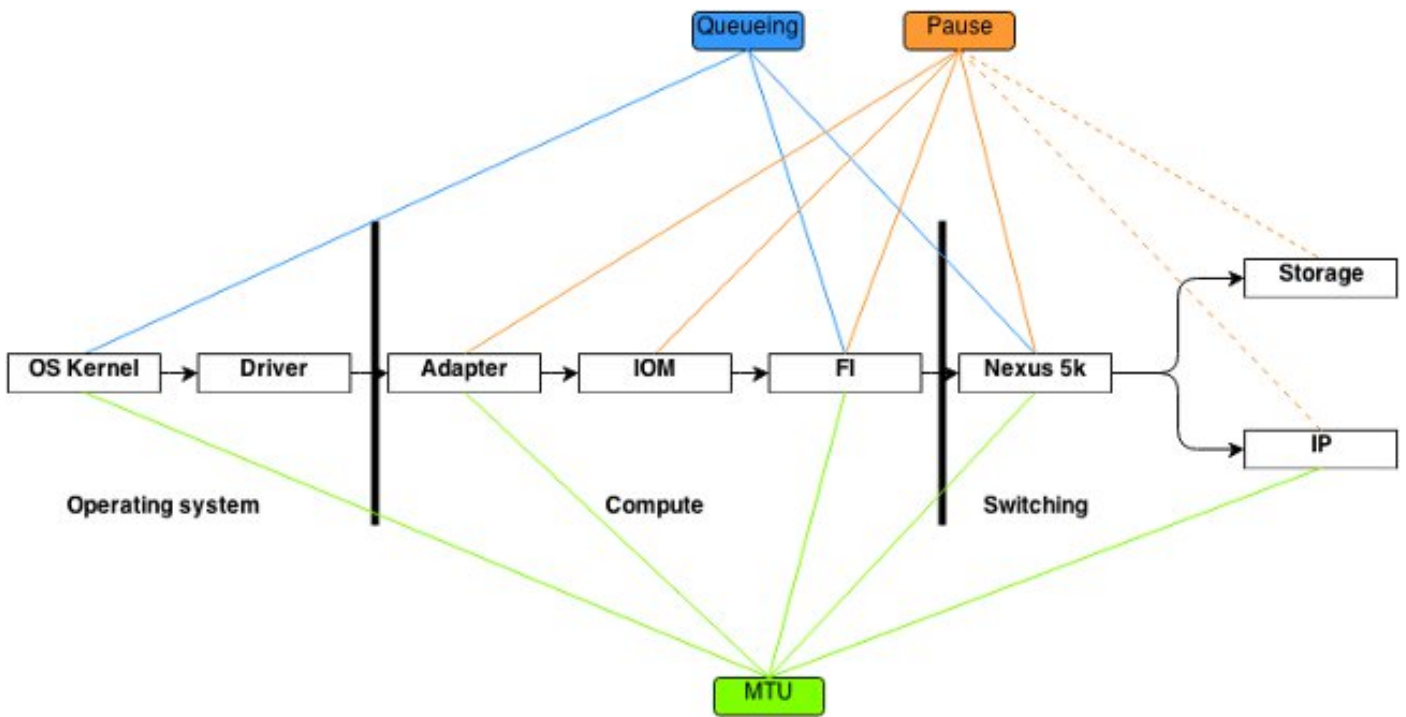
 784.185kbps                Rx rate

```

两有趣位信息表示，由适配器由于缺乏缓冲区或缓冲被禁用和另外TCP数据段丢弃卸载(TSO)分段被处理的96个帧。

逻辑信息包流

显示的图表这里在FlexPod环境里概述逻辑信息包流。



此图表意味着，帧在途中通过通过FlexPod环境组件的细分。它不反射的复杂性任何块并且是方式记住应该配置和验证的地方特定的功能。

输入-输出模块

如逻辑信息包流图表所显示，输入-输出模块(国际移民组织)是组件在通过UCS的所有通信中间。为了连接到在机箱x的IOM，请输入x命令连接的iom。

这是几个其他有用的命令：

- 结构信息-显示平台软件[woodside|红木] sts命令显示拓扑信息从国际移民组织的观点。


```

# show platform software statistics loss

```

Port	SMD			Port Extra Drop	S8 Loss Counters	Cos_q												
	Tx Pkts	Rx Pkts	Errors			Counters	COS											
					Rx PG	Tx PG	CS Total	0	1	2	3	4	5	6	7	8	9	
0-NI2	0	82	0	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	0

由于基础结构工作的方式，计数器为接口仅显示哪些体验两个命令的所有损失介于中间的执行。在本例中，您看到NI2接口接收了82个暂停帧，并且28个暂停帧被传输建立接口HI23，您认识连接前端3。

设计注意事项

FlexPod允许存贮和数据网灵活的配置和设置。使用灵活性也来另外挑战。是重要的跟随最佳实践文件和Cisco被验证的设计(CVD)：

- CVD - [FlexPod部署指南](#)
- NetApp存贮最佳实践(不特定对Flexpod) - [思科统一计算系统\(UCS\)存贮连接选项和最佳实践与NetApp存贮](#)

端口速度选择和端口通道考虑

TAC工程师看到的常见问题是链路的过度使用由于1 Gbit以太网的选择而不是最佳实践文件参考的10 Gbit以太网。作为一个针对性的示例，单流式性能不会是好在于1 Gbit链路与10 Gbit链路比较。在端口通道中单个流可以在单条链路去。

为了欲知什么负载均衡方法在连结和FI的NX-OS使用，请输入show port channel负载均衡命令。管理员能也发现在端口信道建立接口将被选择作为信息包或帧的流出的接口。一个帧的简单的示例在VLAN49的在两台主机之间显示得这里：

```

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

```

存贮特定问题

以前讨论的问题对数据和存储网络是普通。为完整性，性能问题特定对存储区域网络(SAN)也被提及。存贮协议用弹性建立，并且mutli小径仍然被增添。随着技术的出现例如不对称的逻辑单元分配(ALUA)和多重通道的IO (MPIO)，更多灵活性和选项被提交给管理员。

存贮安置

另一个考虑是存贮的安置。FlexPod设计指明存贮将附有在连结交换机。直接地附上存贮不依照CVD。如果最佳实践被跟随，支持与直接地附上存贮的设计。同时，那些设计不严格是FlexPod。

最佳路径选择

这不技术上是Cisco问题，和大多那些选项对Cisco设备透明的。它是选择和坚持的常见问题最佳路径。一个现代设备特有的模块(DSM)可以看到多条路径和需要选择最佳一个此屏幕画面显示四可用路径微软视窗和负载均衡选项的NetApp DSM。

The screenshot shows a table of storage paths and a dialog box for configuring DSM properties. The table lists four paths for Disk0, with their operational and admin states, and initiator information. The dialog box shows the 'Default Load Balance Property' settings, with 'Least Queue Depth' selected.

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...

Data ONTAP(R) DSM Properties

Default Load Balance Property

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

应该根据与存储供应商的一讨论选择推荐的设置。那些设置也许影响性能问题。TAC也许请求您执行的一个典型的测试是一个读/写测试通过仅结构A或结构B。这典型地允许您缩小性能问题到在本文的“常见问题”部分讨论的情况。

VM和Hypervisor流量共享

此点是特定的对估计组件，不管供应商。建立hypervisors的从估计观点是创建两台主机总线适配器(HBAs)，—一个存储网络的简单的方法每个光纤的，并且运行在那两个接口的引导程序LUN数据流和虚拟机存贮数据流。总是推荐分裂引导程序LUN数据流和VM存贮数据流。这允许更好的性能和另外允许在这两的逻辑已分解数据流之间。请参阅“已知问题”部分关于示例。

故障排除提示

缩小问题

和一旦其中任一快速地排除故障，缩小问题和询问合适的问题是非常重要的。

- 哪些设备/applications/VM是(/not)受影响？
- 哪个存贮控制器是(/not)受影响？
- 哪些路径是(/not)受影响？
- 问题(/not)多频繁出现？

Cisco

计数器限制

在本文接口，ASIC队列计数器讨论。计数器也产生一个观点在此刻，因此监控计数器增量是重要的。不可能故意地清除某些计数器。例如，ASIC以前被提及的carmel。

为了提供一个针对性的示例，CRC在接口的出现或丢弃也许不是理想的，但是也许预计他们的值是非零。在转换或初始建立期间，计数器可能在某种程度上，可能上升了。因此注释计数器的增量，并且，当是上次是重要的他们被清除了。

控制层面考虑

当查看计数器时是有用的，知道是重要的某些数据层面问题也许不查找容易的反映到控制层面计数器和工具。作为一个针对性的示例，ethanalyzer是可用的在UCS和连结5000的非常有用的工具。然而，它能只捕获控制层面数据流。数据流捕获是什么TAC经常请求，特别是当不是确切时故障位于的地方。

捕获数据流

在终端主机采取的一个可靠的数据流捕获能显示性能问题的清楚和缩小它相当快速地。连结5000和UCS提供数据流SPAN。特别地，SPANing特定的HBAs的UCS的选项和结构边是有用的。为了得知更多数据流捕获功能，当您监控在UCS时的一次会话，请参阅这些参考：

- [物理和虚拟适配器的\(视频\) UCS流量分析](#)
- [Cisco UCS Manager GUI配置指南-监控数据流](#)

NetApp

NetApp提供完全的一套工具为了排除他们的存贮控制器故障，在他们中是：

- perfstat -一个非常有用的工具，为NetApp维护人员典型地运行
- systat -关于繁忙锉刀如何是，并且什么的提供信息锉刀执行- [NetApp支持库](#)

有在最普通的命令中：

- ```
sysstat -x 2
```

- ```
sysstat -M 2
```

这是寻找的一些事在sysstat -也许指示被超载的NetApp阵列或磁盘2输出的x：

- 与大量的持续的CPU列：或者F
- 在20%上的持续的HDD util列

此条款描述如何配置NetApp：[NetApp以太网存贮最佳实践](#)。

- VLAN标记
- VLAN中继
- 超大MTU
- IP散列
- 功能失效流控制

VMware

ESXi提供安全壳SSH访问，您能排除故障。在有用的工具中提供给管理员是esxtop和perfmon。

- esxtop -很象Linux/BSD顶层，允许用户监控实时与表现有关的参数
[使用识别的esxtop存贮ESX/ESXi的性能问题](#)
- perfmon -允许用户排除微软视窗虚拟机(VM)故障
[收集Windows Perfmon日志数据诊断虚拟机性能问题](#)
- 收集在ESXi的的诊断的套件-[收集VMware ESX/ESXi诊断信息使用vSphere客户端\(653\)](#)
- VMware vSwitch Cisco B系列服务器的负载均衡需求-[根据IP哈希的路由没有用Cisco UCS B200 M1/M2使用UCS 6100系列结构互联的前端服务器支持](#)

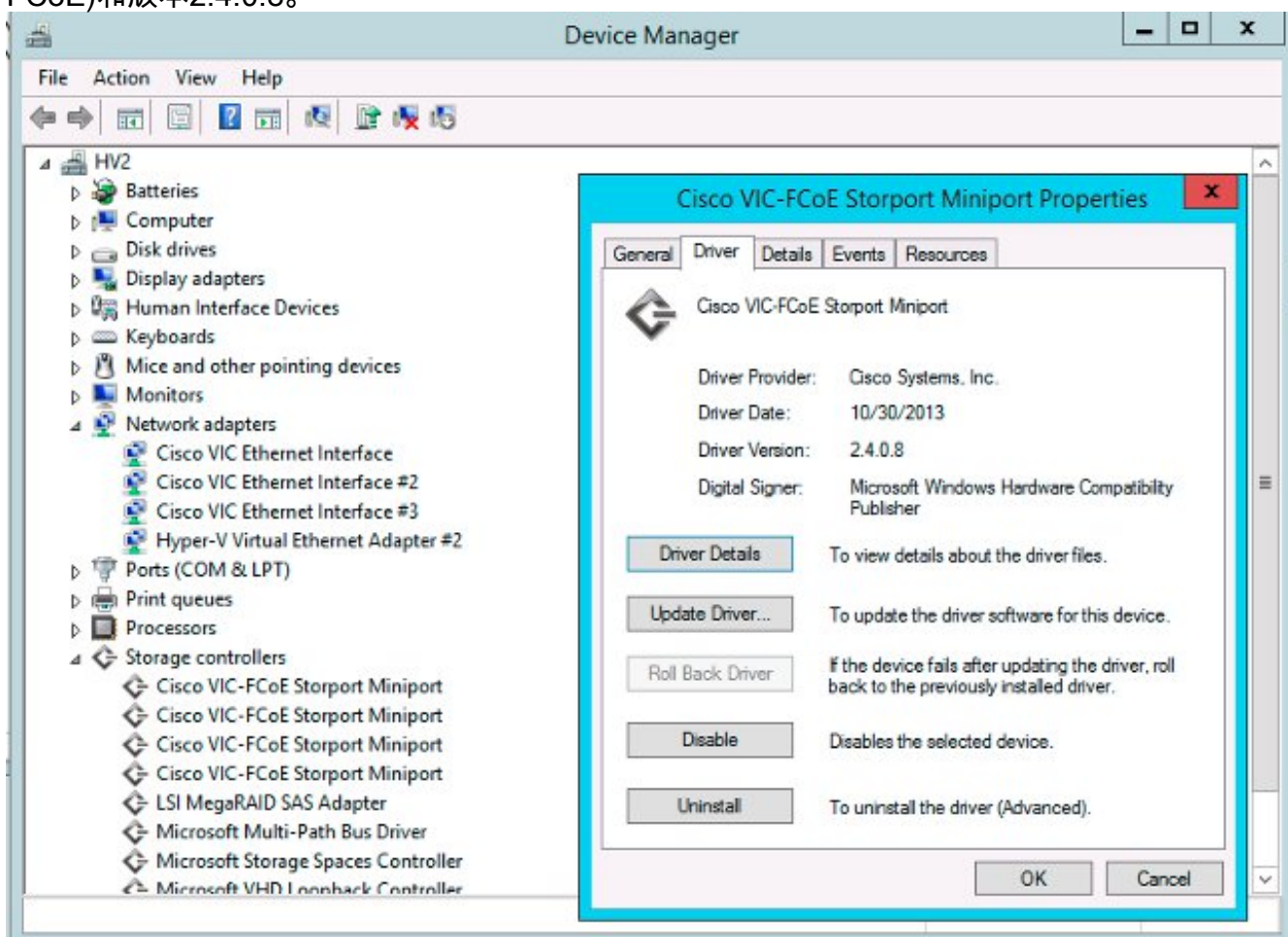
已知问题和增进

- Cisco Bug ID [CSCuj86736](#) -与被动twinax电缆CRC错误可以增加。当连结5000不优化DFE时，这导致。输入**show hardware内部carmel眼睛**命令为了验证“眼睛高度”参数在100 mv上。这在版本被修正了5.2(1)N1(7)和7.0(4)N1(1)。
- Cisco Bug ID [CSCuo76425](#) -类似于早先Bug并且在UCS结构存在互联。这在版本2.2(3a)被修正。
- Cisco Bug ID [CSCuo76425](#) -同一样烦扰除了UCS结构互连的[CSCuj86736](#)的。
- Cisco Bug ID [CSCup40056](#) -共享引起的计时问题引导程序数据流与在[统一计算系统虚拟机实际迁移](#)描述的VM数据流用虚拟光纤信道适配器失效。
- 缓慢的流失检测和避免-经常FC和FCoE是受慢的流失的影响的。NX-OS版本7.0(0)N1(1)引入平均值发现和避免它。得知更多在[Cisco连结5500系列NX-OS接口配置指南](#)的功能并且[减慢流失设备检测和拥塞避免](#)。
- Cisco Bug ID [CSCuj81245](#) -限制存在于PALO该基于的卡(VIC1240和其他)原因FC中止。
- Cisco Bug ID [CSCuh61202](#) -，在发布2.1(3)的升级，UCS固件FC中止后，并且多个其他问题能被看到。
- Cisco Bug ID [CSCtw91018](#) - MTU设置的混合VNICs的在单个，基于PALO的适配器能导致一些缺乏话务类别。
- Cisco Bug ID [CSCuq40256](#) -在从结构互连的链路将造成PFC被禁用下来服务器适配器。这将导致从光纤信道中止开始，并且关于存贮报告的故障中帧支持问题的种类。存贮断开，并且其他性能问题也许被报告。

TAC案例

在许多，在调查可以开始前，案件，TAC工程师将要求您收集一些基本信息。

- 包括端口号和线路速度的拓扑图-，绝对必要。
- UCSM技术支持-[收集技术支持文件的视觉指南\(B和C系列\)](#)。
- UCS体验的一个机箱的机箱技术支持问题-请参阅早先链路。
- 两连结5000技术支持和任何其他网络设备在UCS和NetApp之间-[重定向show tech-support详细资料的输出请发出命令](#)。
- 输出的show queueing interface命令在两个FIs。
sysstat -M 2
- 在ESXi的主机驱动器版本实行-请输入这些命令：vmkload_mod - enic的svmkload_mod - fnic的s
- Linux -
sysstat -M 2
- Windows -请检查在“设备管理器的”驱动版本。从窗口2012 R2的一个示例显示三个Cisco VIC以太网接口和四个VIC FCoE miniport fnic驱动程序的接口(负责也对光纤信道，不仅FCoE)和版本2.4.0.8。



反馈

请使用feedback按钮提供关于本文或您的经验的反馈。我们在反馈以后不断地将更新本文，发展发生和被接受。