

# Problèmes de performances de terrain communal de FlexPod

## Contenu

[Introduction](#)

[Aperçu conceptuel de FlexPod](#)

[Considérations de représentation](#)

[Environnement](#)

[Mesure](#)

[Spécification de base](#)

[Problèmes de performances dans un FlexPod](#)

**[Problèmes courants](#)**

[Trame et perte de paquets](#)

[Non-concordance de MTU](#)

[Affichage de MTU sur le Nexus 5000 et les Plateformes UCS](#)

[Configuration de bout en bout](#)

[Trames étendues de bout en bout de test](#)

[Problèmes relatifs de mémoire tampon](#)

[Problème de pilote](#)

[Les informations d'adaptateur](#)

[Écoulement logique de paquet](#)

[Module d'entrée/sortie](#)

[Considérations de conception](#)

[Considérations de sélection et de Port canalisé de vitesse du port](#)

[Problèmes de particularité de mémoire](#)

[Placement de mémoire](#)

[Sélection de chemin optimal](#)

[Partager du trafic VM et de Hypervisor](#)

**[Conseils de dépannage](#)**

[Rétrécissez vers le bas le problème](#)

[Cisco](#)

[Contre- limites](#)

[Contrôlez les considérations plates](#)

[Le trafic de capture](#)

[NetApp](#)

[VMware](#)

[Problèmes connus et améliorations](#)

[Cas TAC](#)

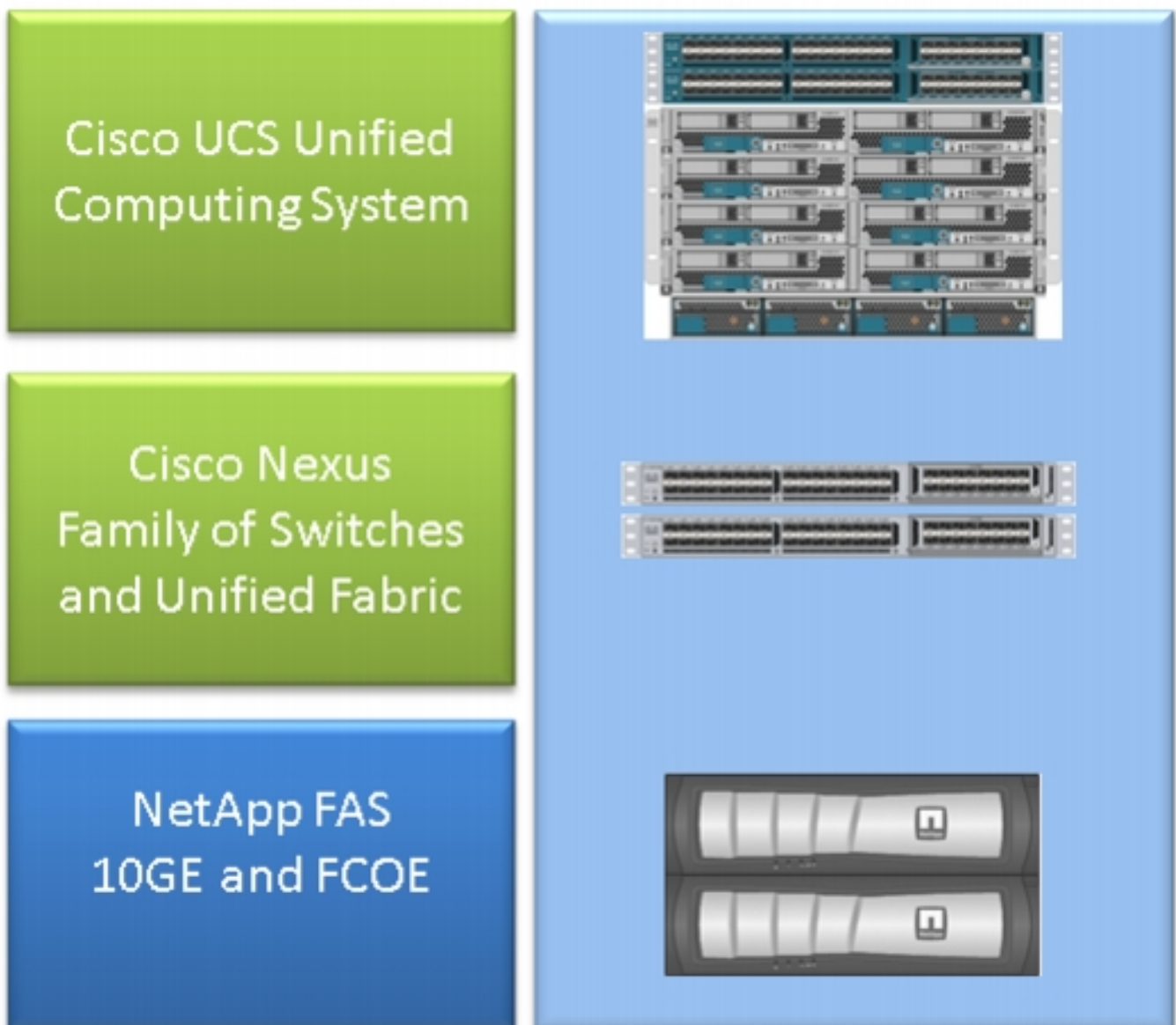
[Commentaires](#)

# Introduction

Ce document décrit des problèmes de performances communs dans des environnements de FlexPod, fournit une méthode pour dépanner des questions, et fournit des étapes de réduction. On le destine en tant que commencer le point pour les clients qui regardent pour dépanner la représentation dans un environnement de FlexPod. Ce document a été écrit en raison des questions vues par l'équipe du centre d'assistance technique de solutions de Data Center (TAC) ces derniers mois.

## Aperçu conceptuel de FlexPod

Un FlexPod se compose d'un ordinateur connecté de l'Unified Computing System (UCS) par l'intermédiaire d'un commutateur de Nexus à la mémoire et aux réseaux IP de NetApp.



Le FlexPod le plus commun se compose de la B-gamme d'un Cisco UCS que le châssis connecté par l'intermédiaire de la matrice interconnectée (FIs) aux Commutateurs du Nexus 5500 aux fichiers de NetApp. Une autre solution, appelée le FlexPod exprès, utilise un châssis de série C UCS connecté aux Commutateurs du Nexus 3000. Ce document discute le FlexPod le plus commun.

# Considérations de représentation

Dans les environnements complexes avec les interlocuteurs responsables de multiple comme typiquement vus dans un FlexPod, vous devez considérer de plusieurs aspects afin de dépanner la question. Les problèmes de performances typiques dans la couche 2 et les réseaux IP proviendraient :

- Paquet ou perte de trame - la perte de bits de données entraîne un effet inverse sur la représentation des applications.
- Bufferisant - si un paquet ou une trame passe trop d'heure dans une file d'attente/mémoire tampon que certaines implications de représentation pourraient être vues par des applications, particulièrement en cas de Réseaux de stockage. La latence, le réarrangement, et les problèmes de normalisateur tombent sous cette catégorie.
- Problèmes et fragmentation de non-concordance de MTU - un problème courant quand vous atteignez la performance supérieure. Questions qui associent à la chute de fragmentation et d'incohérence de MTU dans cette catégorie.

## Environnement

Il est important de connaître l'environnement le quel la représentation est mesurée. Les questions au sujet de la mémoire tapent et protocole, aussi bien que le système d'exploitation du serveur affecté (SYSTÈME D'EXPLOITATION) et l'emplacement, devraient être augmentés pour rétrécir correctement le problème vers le bas. Un diagramme de topologie qui trace les grandes lignes de la Connectivité est le strict minimum.

## Mesure

Vous devez savoir ce qui est mesuré et comment il est mesuré. Certaines applications, aussi bien que la plupart des constructeurs de mémoire et de hypervisor, fournissent les mesures d'un certain tri qui indiquent la représentation/santés du système. Ces mesures sont un point positif à commencer à car elles ne sont pas une substitution pour la plupart des méthodologies de dépannage.

Comme exemple, une mesure de latence de mémoire de Systèmes de fichiers en réseau (NFS) dans le hypervisor pourrait indiquer que la représentation descend, toutefois seule elle n'implique pas le réseau. Dans le cas d'un NFS, un ping simple de l'hôte au réseau IP de mémoire NFS pourrait indiquer si le réseau est de blâmer.

## Spécification de base

Ce point ne peut pas être soumis à une contrainte assez, particulièrement quand vous ouvrez une valise TAC. Afin d'indiquer que la représentation est insatisfaisante, le paramètre mesuré doit être indiqué. Ceci inclut la valeur prévue **et** testée. Dans le meilleur des cas, vous devriez afficher des données précédentes **et la** méthodologie de test utilisée pour réaliser ces données.

Comme exemple ; la latence 10ms réalisée une fois testée, avec un écriture seulement d'un demandeur simple à un seul numéro d'unité logique (LUN), ne pourrait pas être indicative de ce

que la latence est censée pour être pour un système entièrement chargé.

## Problèmes de performances dans un FlexPod

Puisque ce document est destiné comme référence pour la majorité d'environnements de FlexPod, il trace les grandes lignes seulement des problèmes les plus fréquents comme vu par l'équipe TAC responsable des solutions de Data Center.

### [Problèmes courants](#)

Des problèmes communs à la mémoire et les réseaux IP/Layer 2 sont discutés dans cette section.

#### Trame et perte de paquets

La vue et la perte de paquets est le facteur le plus fréquent qui affecte la représentation. Un des endroits communs pour rechercher des indications d'un problème est au niveau d'interface. Du Nexus 5000 ou du système d'exploitation de Nexus UCS (NX-OS) CLI, écrivez l'**interface d'exposition | sec « est en hausse » | ^ d'egrep (Eth|fc)|jetez|baisse|Commande de CRC**. Pour les interfaces qui sont en hausse, il répertorie le nom et jette des compteurs et des baisses. De même, un grand aperçu est affiché quand vous sélectionnez la commande d'**erreur de compteurs d'interface d'exposition** qui affiche des statistiques sur les erreurs pour toutes les interfaces.

#### Monde d'Ethernets

Il est important de savoir que les compteurs à non-0 ne pourraient pas indiquer un problème. Dans certains scénarios ces compteurs pourraient avoir été augmentés dans la première installation ou dans les modifications opérationnelles précédentes. Une augmentation des compteurs devrait être surveillée.

On peut également recueillir des compteurs du niveau ASIC, qui pourrait être plus indicatif. Spécifiquement, pour l'erreur de contrôle de redondance cyclique (CRC) sur des interfaces, une commande préférée TAC d'entrer est **crc interne de carmel de matériel d'exposition**. Carmel est le nom de l'ASIC responsable de l'expédition niveau du port.

La sortie semblable peut être prise de la gamme 6100 FIs ou des Commutateurs du Nexus 5600 sur une base de par-port. Pour le fi 6100, les gatos ASIC, sélectionnent cette commande :

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

Pour le Nexus 5600, du bigsur ASIC, sélectionnez cette commande :

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

La commande pour le carmel ASIC affiche à où des paquets de CRC ont été reçus et à où ils ont été expédiés, et d'une manière primordiale, qu'ils aient été frappés du pied ou pas.

Puisque le Nexus 5000 et l'exécution UCS NX-OS est cut-through, des trames de changement de mode avec le Frame Check Sequence incorrect (FCS) sont seulement frappées du pied avant la

transmission. Il est important de découvrir où les trames corrompues proviennent.

```
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 |
```

Cet exemple affiche les paquets frappés du pied qu'Eth provenu 1/17 et Eth 1/18, qui est une liaison ascendante au Nexus 5000. On peut supposer que ces trames ont été plus tard envoyées vers le bas à Eth 1/34, tel qu'Eth 1/17 + Eth que 1/18 rx frappent du pied = Eth 1/34 tx frappent du pied.

Un aspect semblable sur les expositions de Nexus 5000 :

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

Cette sortie affiche que des crc reçus sur deux liens et marqués en tant que frappe du pied avant la transmission. Le pour en savoir plus, voyent le [guide de dépannage de Nexus 5000](#).

## Monde de la Manche de fibre

Un moyen simple de rechercher des baisses (discrds, erreur, crc, épuisement de crédit de commerce électronique interentreprises) est par l'intermédiaire de la commande de **fc de compteurs d'interface d'exposition**.

Ces commande, disponibles sur le Nexus 5000 et le Fabric Interconnect, donne une bonne indication de ce qui se produit dans le monde de la Manche de fibre.

Exemple :

```
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
(...)
```

Cette interface n'est pas occupée, et la sortie prouve qu'aucun écart ou erreur ne s'est produit.

Supplémentaire, des transitions de crédit de commerce électronique interentreprises de 0 ont été mises en valeur ; en raison des id [CSCue80063](#) et [CSCut08353 de](#) bogues Cisco, ces compteurs ne peuvent pas s'appuyer sur la confiance. Ils fonctionnent bien sur le Cisco MDS, mais pas sur l'UCS des Plateformes Nexus5k. Également vous pouvez vérifier l'ID de bogue Cisco [CSCsz95889](#).

De même au Carmel en mode d'Ethernets pour la Manche de fibre (FC) l'installation de fc-MAC peut être utilisée. Comme exemple, pour le port fc2/1, sélectionnez la commande **interne de statistiques du port 1 du fc-MAC 2 de matériel d'exposition**. Les compteurs présentés sont dans le format hexadécimal.

```
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
```

La sortie affiche 15 écarts sur l'entrée. Ceci peut être apparié à FCP\_CNTR\_PIF\_RX\_DROP qui a compté à 0xf (15 dans la décimale). Ces informations peuvent être de nouveau corrélées aux informations FWM (gestionnaire d'expédition).

```
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
```

Cependant, ce tellis l'administrateur la quantité de baisses et qui est le nombre correspondant ASIC. Les informations d'obtenir sur la raison de cela ASIC abandonné doivent être questionnées.

```
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]
```

Dans ce cas, le trafic a été abandonné par la liste de contrôle d'accès d'entrée (ACL), typiquement en mode FC - Répartition en zones.

## Non-concordance de MTU

Dans des environnements de FlexPod il est important de faciliter la configuration maximum de bout en bout d'unité de transition (MTU) pour des applications et des protocoles où on l'exige. Dans le cas de la plupart des environnements, c'est la Manche de fibre au-dessus des Ethernets (FCoE) et des Trames étendues.

Supplémentaire, la fragmentation se produit, la représentation dégradée doit être prévue. En cas de protocoles tels que le Systèmes de fichiers en réseau (NFS) et l'interface SCSI d'Internet (iSCSI), il est important de tester et prouver la taille maximum de bout en bout d'unité de transmission maximale d'IP (MTU) et de segment de TCP (MSS).

Si vous dépannez des Trames étendues ou FCoE, il est important de se souvenir que chacun des deux ceux ont besoin de configuration cohérente et de Classe de service (Cos) marquant à travers l'environnement afin de fonctionner correctement.

Dans le cas de l'UCS et du Nexus, une commande qui est utile pour valider la par-interface, par configuration de MTU de QoS-groupe est **show queueing interface | i s'alignant|qos-groupe|MTU**.

## Affichage de MTU sur le Nexus 5000 et les Plateformes UCS

Un aspect connu d'UCS et de Nexus est l'affichage des mtu sur l'interface. Cette sortie explique une interface configurée pour aligner des Trames étendues et FCoE :

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

En même temps, la **commande d'interface d'exposition** affiche 1500 octets :

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

Si comparé aux informations du carmel ASIC, l'ASIC affiche la capacité de MTU d'un port donné.

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

Cette non-concordance de MTU dans l'affichage est prévue sur les Plateformes mentionnées ci-dessus, et pourrait potentiellement tromper des débutants.

## Configuration de bout en bout

La configuration cohérente de bout en bout est la seule manière de garantir la représentation appropriée. Des Trames étendues configuration et étapes pour le côté de Cisco, aussi bien que le VMware ESXi, sont décrits dans l'[UCS avec l'exemple enorme de bout en bout de configuration de MTU d'ESXi de VMware](#).

L'[exemple de configuration de liaison ascendante UCS FCoE](#) affiche une configuration UCS et de Nexus 5000. Voir l'annexe A dans le document de référence pour un contour d'une configuration de base de Nexus 5000.

[Installez la Connectivité de FCoE pour des foyers d'une lame de Cisco UCS](#) sur la configuration UCS pour FCoE. [Le Nexus 5000 NPIV FCoE avec FCoE NPV a relié des foyers d'exemple de configuration UCS](#) sur la configuration de Nexus.

## Trames étendues de bout en bout de test

La plupart des systèmes d'exploitation modernes de jour offrent la capacité de tester une configuration appropriée de Trames étendues avec un test simple de Protocole ICMP (Internet Control Message Protocol).

### Calcul

9000 octets - En-tête IP sans options (20 octets) - en-tête d'ICMP (8 octets) = 8972 octets de données

### Commandes dans des systèmes d'exploitation populaires

#### 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
```

## Problèmes relatifs de mémoire tampon

Problèmes associés par latence bufferisants et autres sont parmi les causes communes de dégradation de représentation dans l'environnement de FlexPod. Non tous les problèmes signalés comme latence proviennent des problèmes réels de mise en mémoire tampon, tout à fait quelques mesures pourraient indiquer la latence de bout en bout. Par exemple, dans le cas de NFS, le délai prévu signalé pourrait être avec succès lecture/écriture nécessaire à la mémoire et à la latence de réseau non réelle.

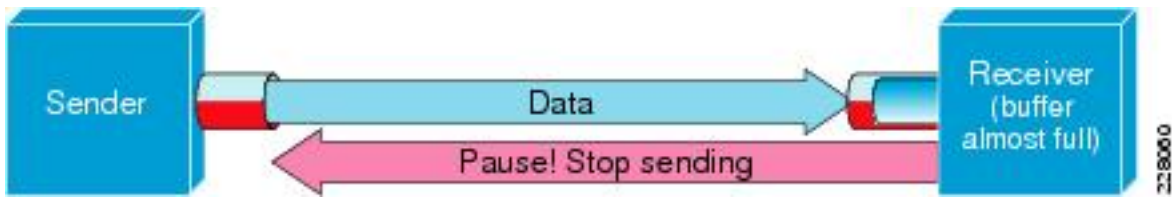
L'encombrement est la plupart de cause classique pour bufferiser. Dans le monde de la couche 2, l'encombrement peut entraîner la mise en mémoire tampon et suit même des baisses des trames. Afin d'éviter des baisses au cours des périodes d'encombrement, les trames de pause d'IEEE 802.3x et le contrôle de flux prioritaire (PFC) ont été introduits. Chacun des deux se fondent sur demander au point final pour tenir des transmissions pendant une courte période tandis que l'encombrement dure. Ceci peut être provoqué par un encombrement de réseau (accablez reçu avec la quantité de données) ou parce qu'une trame prioritaire doit passer, comme dans le point de droit pour FCoE.

### Contrôle de flux - 802.3x

Afin de vérifier que les interfaces ont un contrôle de flux activé, sélectionnez la commande de **flowcontrol d'interface d'exposition**. Il est important de suivre la recommandation du constructeur de mémoire en vue de le contrôle de flux étant activé.

Une illustration qui affiche comment des travaux du contrôle de flux 802.3x sont affichés ici.



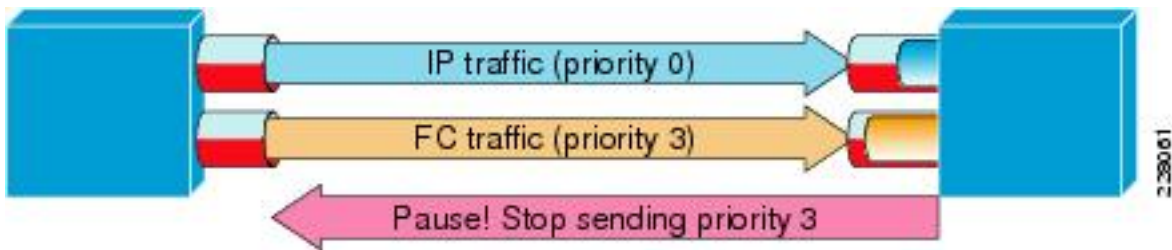


## PFC - 802.1Qbb

Le PFC n'est pas exigé pour toutes les installations, mais est recommandé pour les la plupart. Afin de vérifier que les interfaces ont le PFC activé, le priorité-écoulement-**control d'interface d'exposition** | i sur la commande peut être exécuté sur le NX-OS et le Nexus 5000 de l'UCS.

Les interfaces entre FIs et le Nexus 5000 devraient être visibles sur cette liste. Sinon, la configuration QoS doit être vérifiée. QoS doit être à de bout en bout cohérent afin de tirer profit du PFC. Afin de vérifier pourquoi le PFC ne monte pas sur une interface spécifique, sélectionnez la commande **interne des Ethernets x/y d'interface de log de dcbx de show system** afin d'obtenir Data Center jetant un pont sur le log du protocole d'échange de capacités (DCBX).

Une illustration que des expositions comment la pause encadre travail avec le PFC est affiché ici.



La commande de priorité-écoulement-**control d'interface d'exposition** permet à l'administrateur pour observer par-QoS comportement de classe des trames de pause prioritaire.

Voici un exemple :

```
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)
```

Cette sortie prouve que, dans la deuxième classe, le périphérique transmettait juste (TX) une trame PPP.

Dans ce cas, l'Ethernet 1/1 est port faisant face à IOM et alors que le port global n'aura pas le PFC activé, il pourrait traiter des trames PPP pour des ports 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
```

Dans ce cas, les interfaces FEX sont impliquées.

```
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
```

Les ports FEX qui sont impliqués peuvent être également vérifiés par l'intermédiaire du **détail du flex X d'exposition** où X est le numéro de châssis.

```
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
```

Voir les ces documents pour plus d'informations sur des mécanismes de pause.

- [La Manche de fibre au-dessus des fonctionnements d'Ethernet](#)
- [La Manche blanche de Papier-fibre de structure unifiée au-dessus des Ethernets \(FCoE\)](#)

## Écarts de queue

Le Nexus 5000 et l'UCS NX-OS maintiennent des écarts d'entrée dus à la queue sur a par base de QOS-groupe. Exemple :

```
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)
```

L'écart d'entrée *devrait* se produire seulement dans les files d'attente qui sont configurées pour permettre des baisses.

Les écarts de queue d'entrée peuvent se produire en raison de ces raisons :

- Session du Fonction Switched Port Analyzer (SPAN) /Monitoring activée sur certaines des interfaces (voir l'ID de bogue Cisco [CSCur25521](#))
- De la contre-pression d'une autre interface, des trames de pause sont typiquement vues une fois activée
- Le trafic donné un coup de volée à la CPU

## Problème de pilote

Cisco fournit deux gestionnaires du système d'exploitation pour l'UCS, enic et fnic. Enic est responsable de la connectivité Ethernet et fnic est responsable de la Connectivité de la Manche et de FCoE de fibre. Il est **très important** que les gestionnaires enic et fnic soient exactement comme spécifiés dans la [matrice d'Interopérabilité UCS](#). Les problèmes introduits par les gestionnaires incorrects s'étendent de la perte de paquets et de la latence ajoutée à un plus long processus de démarrage ou absence totale de Connectivité.

## Les informations d'adaptateur

Cisco-a fourni l'adaptateur peut fournir une bonne mesure au sujet du trafic qui est passé, aussi bien que relâche. Cet exemple affiche comment se connecter au châssis X, au serveur Y, et à l'adaptateur 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.
```

D'ici, l'administrateur peut ouvrir une session au centre de surveillance pour l'installation de la représentation (MCP).

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

L'installation MCP te permet pour surveiller l'utilisation du trafic par interface logique (LIF).

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

```
-----
id name          v n i c          l i f          v i f
   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
```

Les châssis 1, divisent 1, et l'adaptateur 1 ont deux networks interface cards virtuels (VNICs) associés avec des interfaces virtuelles (les Ethernets virtuels ou la Manche virtuelle de fibre) 834 et 836. Ceux ont les numéros 2 et 3. Les statistiques pour LIF 2 et 3 peuvent être vérifiées comme affiché ici :

```
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

```

Il est important de noter que l'administrateur de l'UCS est équipé de colonnes de total et de delta (entre deux exécutions ultérieures des lifstats) aussi bien que de charge de la circulation en cours par-LIF et d'informations sur toutes les erreurs qui pourraient s'être produites.

L'exemple précédent n'affiche des interfaces sans aucune erreur avec un chargement très petit. Cet exemple affiche un serveur différent.

```

adapater 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

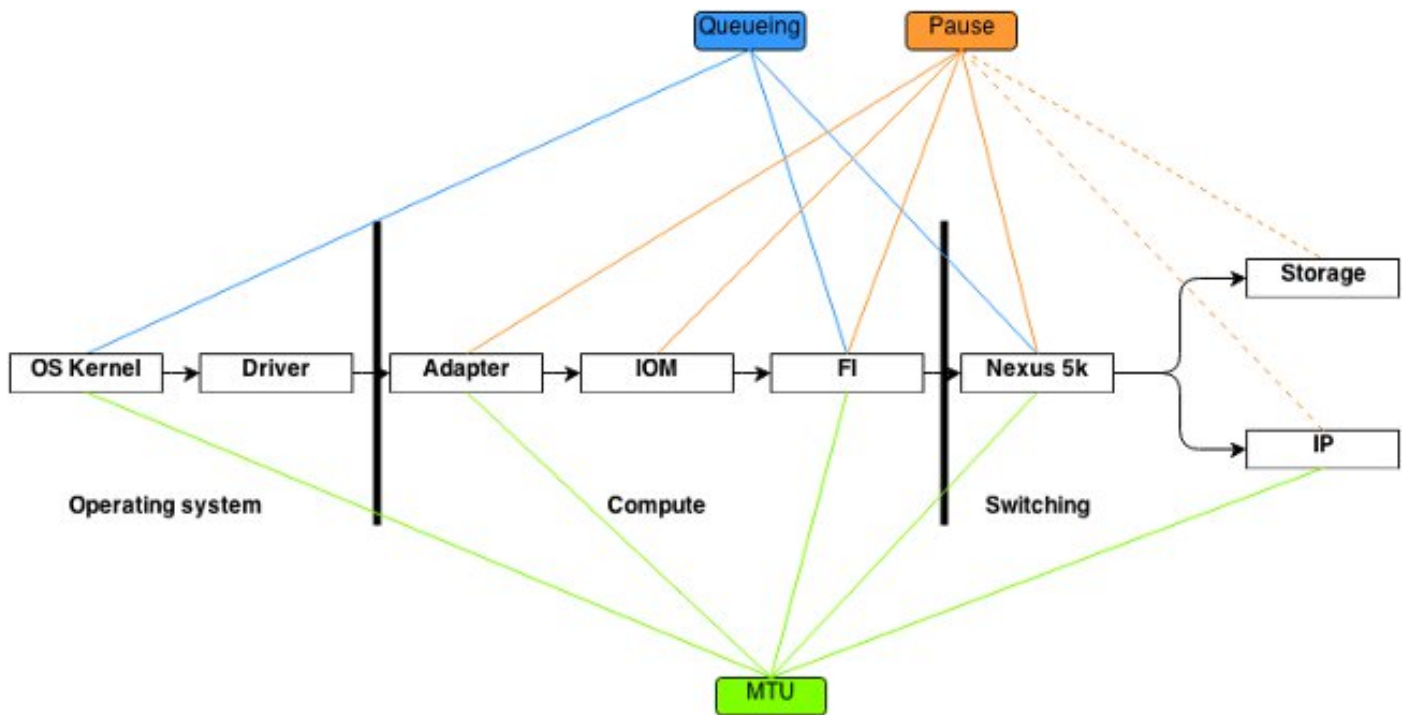
784.185kbps          Rx rate

```

Deux bits intéressants des informations prouvent que 96 trames ont été abandonnées par l'adaptateur devant manquer de la mémoire tampon ou la mise en mémoire tampon a désactivé et supplémentaire segment de débarquement de segment de TCP (TSO) étant traitée.

## Écoulement logique de paquet

Le diagramme affiché ici trace les grandes lignes de l'écoulement logique de paquet dans un environnement de FlexPod.



Ce diagramme est signifié comme une panne des composants qu'une trame a traversés sur le chemin par l'intermédiaire de l'environnement de FlexPod. Il ne reflète pas la complexité des blocs l'uns des et est simplement une manière de mémoriser où des fonctions particulières devraient être configurées et vérifiées.

## Module d'entrée/sortie

Suivant les indications de l'organigramme logique de paquet, le module d'entrée/sortie (IOM) est un composant au milieu de toute la transmission qui passe par l'UCS. Afin de se connecter à l'IOM dans des châssis X, sélectionnez la commande de l'**iom X de connecter**.

Voici plusieurs autres commandes utiles :

- Les informations topologiques - le **logiciel de show platform [woodside|le sts de séquoia]** commande les informations topologiques d'expositions du point de vue de l'IOM.



```

# show platform software sidekick loss

```

Port	SMD			Port Extra Drop Counters	S8 Loss Counters	Can_u															
	Tx Pkts	Rx Pkts	Errors			Rx PG	Tx PG	S8 Total	COS												
									0	1	2	3	4	5	6	7	8	9	10	11	
0-NI2	0	82	0	0	0	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	0	0	0

En raison de la manière que l'infrastructure sous-jacente fonctionne, les compteurs sont affichés seulement pour des interfaces ce qui ont éprouvé n'importe quelle exécution intermédiaire de perte des deux commandes. Dans cet exemple, vous voyez que l'interface NI2 a reçu 82 trames de pause et que 28 trames de pause ont été transmises pour relier HI23, que vous connaissez est relié à la lame 3.

### Considérations de conception

Un FlexPod tient compte d'une configuration souple et d'une installation de mémoire et de mise en réseau des données. Avec la flexibilité est livré également des défis supplémentaires. Il est essentiel de suivre les documents de pratiques recommandées et une conception validée par Cisco (CVD) :

- CVD - [Guide de déploiement de FlexPod](#)
- Pratiques recommandées de mémoire de NetApp (non spécifiques à Flexpod) - [options et pratiques recommandées de Connectivité de mémoire du Système d'informatique unifiée Cisco \(UCS\) avec la mémoire de NetApp](#)

### Considérations de sélection et de Port canalisé de vitesse du port

Un problème courant vu par des ingénieurs TAC est overutilization des liens dus à la sélection de 1 Ethernet de Gbit au lieu de 10 Ethernets de Gbit référencés dans des documents de pratique recommandée. Comme exemple aigu, la représentation à **courant simple** ne sera pas meilleure sur dix liens de 1 Gbit comparés à un lien de 10 Gbit. Dans le Port canalisé un à courant simple peut aller au-dessus d'un lien simple.

Afin de découvrir quelle méthode d'Équilibrage de charge est utilisée sur NX-OS de Nexus et/ou de fi, sélectionnez la commande de **show port-channel load-balance**. L'administrateur peut également découvrir qui relie dans un Port canalisé seront choisis comme interface sortante pour un paquet ou une trame. Un exemple simple d'une trame sur VLAN49 entre deux hôtes est affiché ici :

```

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

```

## Problèmes de particularité de mémoire

Les problèmes discutés précédemment sont communs aux données et aux Réseaux de stockage. Dans l'intérêt de l'exhaustivité, des problèmes de performances spécifiques au réseau de stockage (SAN) sont également mentionnés. Des protocoles de mémoire ont été établis avec la résilience et la mutli-voie d'accès sont encore augmentées. Avec l'arrivée des Technologies telles que l'affectation asymétrique d'unité logique (ALUA) et l'E/S par trajets multiples (MPIO), plus de flexibilité et d'options sont présentés aux administrateurs.

## Placement de mémoire

Une autre considération est placement de mémoire. Une conception de FlexPod dicte que la mémoire doit être reliée sur des Commutateurs de Nexus. Directement l'attached storage ne se conforme pas à la CVD. Des conceptions avec directement l'attached storage sont prises en charge, si des pratiques recommandées sont suivies. En même temps, ces conceptions ne sont pas strictement FlexPod.

## Sélection de chemin optimal

Ce n'est pas techniquement Cisco émettent, comme la plupart de ces options sont transparentes aux périphériques de Cisco. C'est un problème courant à sélectionner et coller à un chemin optimal. Un module spécifique de périphérique moderne (DSM) peut être présenté avec des plusieurs chemins et des besoins de sélectionner optimal un celui Ce tir d'écran affiche quatre chemins disponibles à NetApp DSM pour Microsoft Windows et des options d'Équilibrage de charge.

The screenshot shows the 'Data ONTAP(R) DSM Properties' dialog box with the 'MPIO' tab selected. The background window displays a table of paths for Disk0.

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 'Data ONTAP(R) DSM Properties' dialog box shows the following options:

- Default Load Balance Property:
  - Auto Assign
  - Failover Only
  - Round Robin
  - Round Robin with Subset
  - Least Weighted Paths
  - Least Queue Depth

Les configurations recommandées devraient être sélectionnées ont basé sur une discussion avec le constructeur de mémoire. Ces configurations pourraient affecter des problèmes de



performances. Un essai typique au lequel le TAC pourrait te demander de réaliser est un test lecture/écriture par seulement la matrice A ou la matrice B. Ceci te permet typiquement pour rétrécir vers le bas des problèmes de performances aux situations discutées dans la section de « problèmes courants » de ce document.

## **Partager du trafic VM et de Hypervisor**

Ce point est spécifique au composant de calcul, indépendamment du constructeur. Une méthode facile d'installer un réseau de mémoire pour les hypervisors du point de vue de calcul est de créer deux adaptateurs de bus hôte (HBAs), un pour chaque fibre, et exécute chacun des deux le trafic du démarrage LUN et le trafic de la mémoire du virtual machine (VM) au-dessus de ces deux interfaces. Il est toujours recommandé pour séparer le trafic du trafic du démarrage LUN et de mémoire VM. Ceci tient compte d'une meilleure représentation et permet supplémentaire un fractionnement logique entre les deux genres de trafic. Voyez la section de « problèmes connus » pour un exemple.

## **Conseils de dépannage**

### **Rétrécissez vers le bas le problème**

Comme dans le cas de en jeûnent le dépannage, il est très important de rétrécir vers le bas le problème et de poser les questions des droits.

- Quels périphériques/applications/VM sont (/not) affectés ?
- Quel contrôleur de mémoire sont (/not) affecté ?
- Quels chemins sont (/not) affectés ?
- Combien de fois le problème (/not) apparaît-il ?

## **Cisco**

### **Contre- limites**

Dans cette interface de document, des compteurs de Mise en file d'attente ASIC sont discutés. Les compteurs donnent également un avis à un moment, ainsi il est important de surveiller l'augmentation des compteurs. Certains compteurs ne peuvent pas être effacés par conception. Par exemple, le carmel ASIC mentionné précédemment.

Afin de donner un exemple aigu, la présence du CRC ou les écarts sur une interface ne pourrait pas être idéaux, mais il pourrait prévoir que leurs valeurs sont différentes de zéro. Les compteurs pourraient avoir monté à un moment, probablement pendant la transition ou la première installation. Par conséquent il est important de noter l'augmentation des compteurs et quand était la dernière fois ils ont été effacés.

### **Contrôlez les considérations plates**

Tandis qu'il est utile de passer en revue des compteurs, il est important de savoir que certains problèmes de plan de données ne pourraient pas trouver une réflexion facile pour contrôler les compteurs et les outils plats. Comme exemple aigu, l'ethanalyzer est un outil très utile qui est disponible sur l'UCS et le Nexus 5000. Cependant, il peut seulement capturer le trafic d'avion de contrôle. Est une capture du trafic ce que le TAC demande souvent, particulièrement quand il n'est pas clair où le défaut se trouve.

## Le trafic de capture

Une capture digne de confiance du trafic prise sur les hôtes d'extrémité peut jeter la lumière sur un problème de performances et la rétrécir vers le bas tout à fait rapide. Le Nexus 5000 et ENVERGURE du trafic d'offre UCS. Spécifiquement, les options de l'UCS de SPANing HBAs particulier et les côtés de matrice sont utiles. Afin de se renseigner plus sur les capacités de capture du trafic quand vous surveillez une session sur l'UCS, voir les ces références :

- [Analyse du trafic UCS pour les adaptateurs physiques et virtuels](#) (vidéo)
- [Guide de configuration GUI de Cisco UCS Manager - Surveillance du trafic](#)

## NetApp

NetApp offre un ensemble complet d'utilitaires afin de dépanner leurs contrôleurs de mémoire, parmi eux sont :

- perfstat - un utilitaire très utile, fonctionnent typiquement pour le personnel de support de NetApp
- sysstat - fournit des informations au sujet de la façon dont occupé le fichier est et de ce que le fichier fait - [bibliothèque de support de NetApp](#)

Il y a parmi les commandes les plus communes :

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

Voici quelques choses à rechercher dans le `sysstat -x 2` sorti qui pourrait indiquer la baie surchargée ou les disques de NetApp :

- Colonne **ty** soutenue **CP** avec beaucoup de : ou F
- Colonne soutenue **utilisation HDD** au-dessus de 20%

Cet article décrit comment configurer NetApp : [Pratiques recommandées de mémoire d'Ethernets de NetApp](#).

- Étiquetage VLAN
- Jonction VLAN
- MTU énorme
- Hachage IP
- Flowcontrol de débronnement

## VMware

ESXi fournit l'accès de Protocole Secure Shell (SSH), par lequel vous pouvez dépanner. Parmi les la plupart des outils utiles fournis aux administrateurs sont l'esxtop et le perfmon.

- esxtop - tout comme le dessus Linux/BSD, il permet à des utilisateurs de surveiller des paramètres liés à la performance en temps réel  
[Utilisant l'esxtop pour identifier des problèmes de performance de mémoire pour ESX/ESXi](#)
- perfmon - permet à des utilisateurs pour dépanner les virtual machine de Microsoft Windows (la VM)  
[Collecter les données de log de Windows Perfmon pour diagnostiquer des problèmes de performance de virtual machine](#)
- Collectez le paquet diagnostique sur ESXi - [Collectant les informations de diagnostic pour le VMware ESX/ESXi utilisant le client de vSphere \(653\)](#)
- Condition requise d'Équilibrage de charge de vSwitch de VMware pour des serveurs de B-gamme de Cisco - le [routage en fonction sur des informations parasites IP n'est pas pris en charge avec les serveurs lame du Cisco UCS B200 M1/M2 qui utilisent la gamme 6100 UCS que la matrice interconnecte](#)

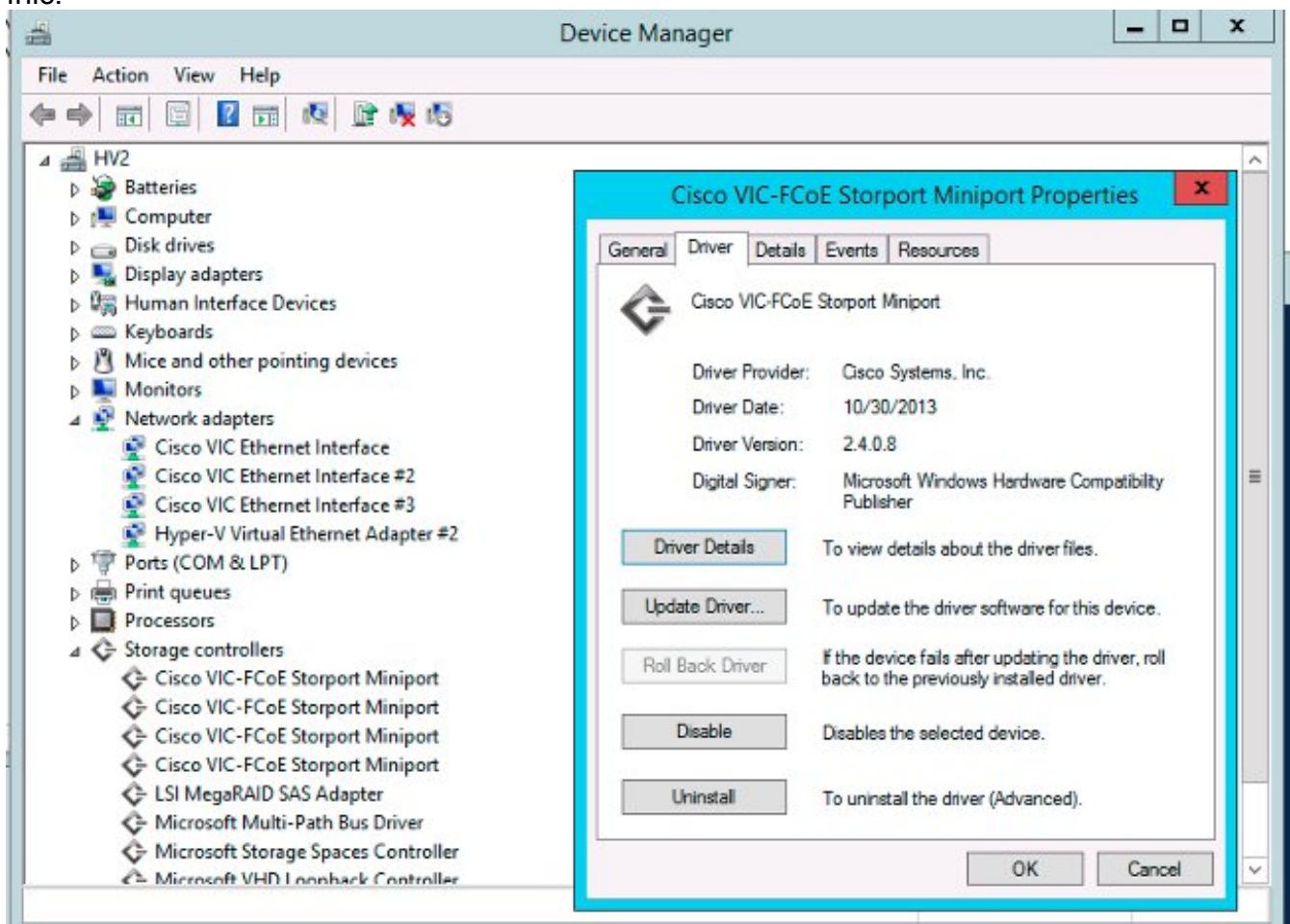
## Problèmes connus et améliorations

- ID de bogue Cisco [CSCuj86736](#) - avec des erreurs passives de CRC de câbles de twinax peut augmenter. Ceci est provoqué par quand le Nexus 5000 n'optimise pas DFE. Sélectionnez la commande **interne d'oeil de carmel de matériel d'exposition** afin de vérifier que le paramètre « de hauteur d'oeil » est au-dessus de 100 système mv. Ceci a été réparé dans des versions 5.2(1)N1(7) et 7.0(4)N1(1).
- L'ID de bogue Cisco [CSCuo76425](#) - semblable à la bogue précédente et existe également sur la matrice UCS interconnecte. Ceci est réparé dans la version 2.2(3a).
- ID de bogue Cisco [CSCuo76425](#) - mêmes que la bogue [CSCuj86736](#) excepté UCS Fabric Interconnect.
- L'ID de bogue Cisco [CSCup40056](#) - problème causé de synchronisation en partageant du trafic de démarrage avec le trafic VM décrit dans le [transfert vivant de virtual machine d'Unified Computing System échoue avec les adaptateurs virtuels de la Manche de fibre](#).
- Détection et manière d'éviter lentes de surcharge - très souvent FC et FCoE sont affectés par la surcharge lente. La version NX-OS 7.0(0)N1(1) introduit des moyens de le détecter et éviter. Renseignez-vous plus sur la caractéristique dans le [guide de configuration d'interfaces de la gamme 5500 NX-OS de Cisco Nexus](#) et [ralentissez la détection de périphérique de surcharge et la manière d'éviter d'encombrement](#).
- ID de bogue Cisco [CSCuj81245](#) - une limite existe dans les cartes basées par PALO (VIC1240 et d'autres) ce des arrêts des causes FC.
- ID de bogue Cisco [CSCuh61202](#) - après que la mise à jour pour libérer 2.1(3), le micrologiciel FC UCS abandonne et le multiple d'autres problèmes peut être vu.
- ID de bogue Cisco [CSCtw91018](#) - un mélange de configurations de MTU pour VNICs sur un adaptateur simple et basé sur PALO peut entraîner la famine pour certaines de classes du trafic.
- L'ID de bogue Cisco [CSCuq40256](#) - causera le PFC d'être désactivé sur des liens de Fabric Interconnect vers le bas aux adaptateurs de serveur. Ceci entraînera la variété de problèmes qui commencent par des arrêts de la Manche de fibre et les trames en panne signalées sur la mémoire dégrossissent. Des débranchements de mémoire et d'autres problèmes de performances pourraient être signalés.

## Cas TAC

Dans plusieurs des cas, l'ingénieur TAC te demandera de collecter quelques informations de base avant qu'une enquête puisse être commencée.

- Diagramme de topologie - qui inclut des numéros de port et des vitesses linéaires, absolument nécessaire.
- Soutien technique UCSM - [Guide visuel pour collecter des fichiers de support technique \(B et série C\)](#).
- Soutien technique de châssis UCS pour un châssis qui éprouve des problèmes - voir le lien précédent.
- Soutien technique de Nexus 5000 et tous autres périphériques de réseau entre l'UCS et le NetApp - [en réorientant la sortie des détails de show tech-support commandez](#).
- Sortie de la commande de **show queueing interface** sur chacun des deux FIs. `sysstat -M 2`
- Les versions de pilote de hôte sur l'ESXi exécutent - sélectionnez ces commandes :  
vmkload\_mod - s enicvmkload\_mod - s fnic
- Linux - `sysstat -M 2`
- Windows - vérifiez la version de gestionnaire dans le « gestionnaire de périphériques ». Un exemple des expositions R2 de la fenêtre 2012 trois interfaces Ethernet de carte d'interface virtuelle de Cisco et quatre interfaces de miniport de FCoE de carte d'interface virtuelle (responsables aussi de la Manche de fibre, non seulement FCoE) et release 2.4.0.8 du gestionnaire fnic.



## [Commentaires](#)

Utilisez le bouton de feedback pour fournir le feedback au sujet de ce document ou de vos

expériences. Nous mettrons à jour continuellement ce document comme les développements se produisent et après feedback est reçu.