

# Dépannage du problème de tous les chemins vers le bas (APD) de Hyperflex Storage NFS

## Contenu

[Introduction](#)

[Comment les data stores HX sont-ils montés sur ESXI ?](#)

[Tous les chemins vers le bas](#)

[Description du problème](#)

[Workflow de dépannage](#)

[Archivage du serveur vCenter :](#)

[Vérifie tous les StCtIVM :](#)

[StCtIVM : StCtIVM d'un hôte ESXi affecté](#)

[Vérifie dans l'hôte ESXi :](#)

## Introduction

Ce document vous donne une compréhension rapide et des étapes de dépannage qui peuvent être effectuées afin d'évaluer la source du problème si vous voyez un message d'erreur « NFS all paths down » dans vCenter avec lequel le cluster Hyperflex est intégré.

## Comment les data stores HX sont-ils montés sur ESXI ?

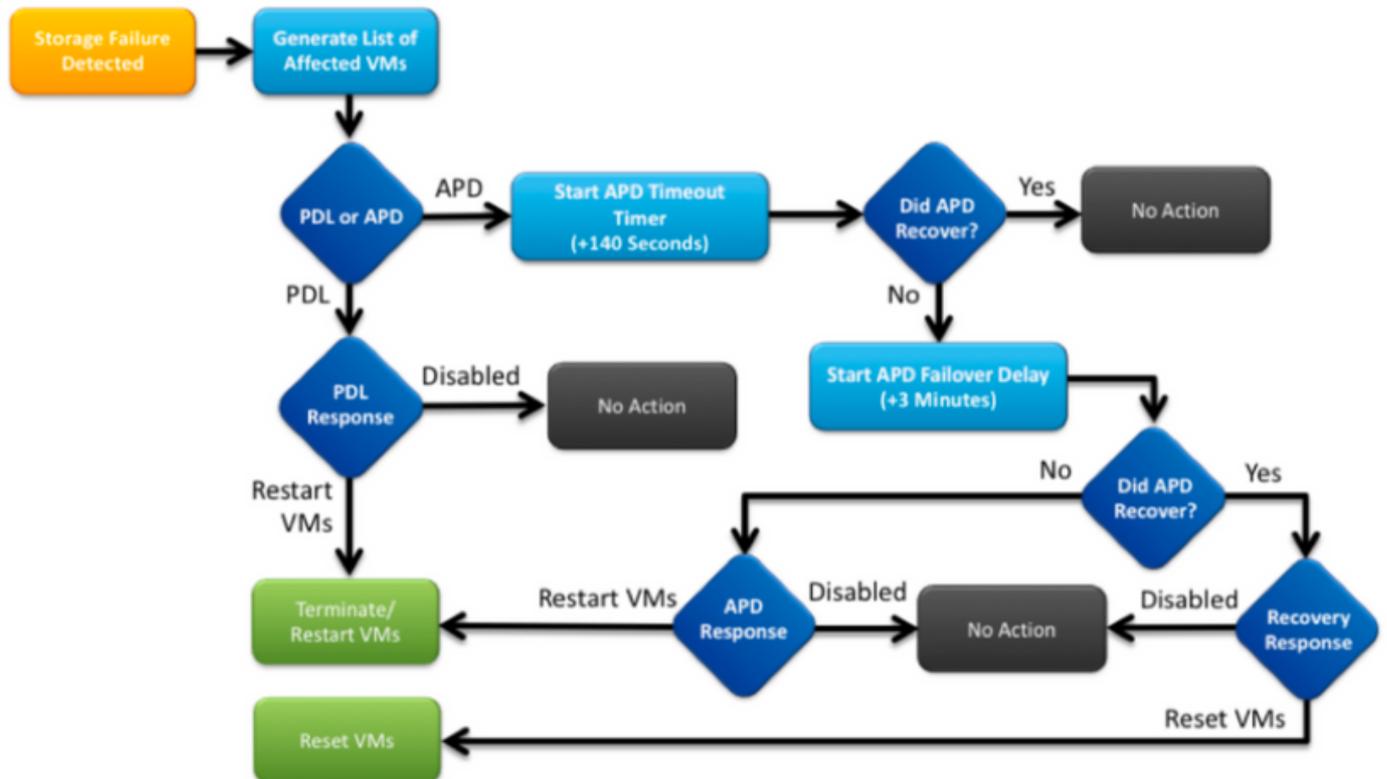
- Les magasins de données hyperflex sont montés sur les hôtes ESXi lors du montage NFS, afin de monter un data store NFS, nous avons besoin de l'IP de serveur NFS qui dans notre cas est l'interface flottante virtuelle eth1:0.
- Le cluster Hyperflex utilise l'utilisation d'**IP flottante** virtuelle à la fois pour la gestion (eth0:mgmtip) et **les données de stockage (eth1:0)** sur lesquelles chaque IP sera attribuée à une machine virtuelle de contrôleur de stockage spécifique (**StCtIVM**). Notez qu'ils peuvent se retrouver dans différents StCtIVM.
- L'importance de ceci est que **l'IP de données de stockage de cluster (eth1:0) est celle utilisée pour monter le magasin de données** créé dans le **cluster Hyperflex**. Il est donc essentiel qu'il soit **assigné et accessible** à partir de tous les **noeuds** du cluster.
- Veuillez noter qu'en cas de défaillance de la machine virtuelle StCtIVM qui possède actuellement une adresse IP virtuelle eth1:0, elle doit « migrer » vers une autre machine virtuelle StCtIVM disponible fonctionnant de la même manière qu'un protocole FHRP (First Hop Redundancy Protocol).

## Tous les chemins vers le bas

- APD signifie que l'hôte ne peut pas atteindre le stockage et qu'aucun code SCSI Perdu de périphérique permanent (PDL) n'est retourné de la baie de stockage.
- Comme il ne sait pas si la perte est temporaire ou non, il continuera à essayer d'établir une communication pour plus de 140 s par défaut (délai) + 3 min (délai de basculement), l'hôte

ESXi commence à échouer tout trafic d'E/S de machine non virtuelle envoyé au périphérique de stockage.

- Pour plus d'informations sur APD dans l'environnement vSphere 6.x, consultez l'article de la base de connaissances VMware [Perte permanente de périphérique \(PDL\) et tous les chemins d'accès \(APD\) dans vSphere 5.x et 6.x](#)
- L'image ci-dessous explique chaque étape intermittente :



Un message d'erreur type dans vCenter est le suivant :

	Status	Name	Defined In
:02.corplex...	Alert	NFS all paths down	SV-VC

## Description du problème

Une fois que vous voyez des alertes APD sur vos hôtes, obtenez les informations ci-dessous pour mieux comprendre la description du problème :

- Si un/plusieurs/tous les hôtes ont été affectés, et si certains hôtes particuliers ont été affectés
- Si des modifications ont été effectuées précédemment (configuration/mise à niveau/etc.)
- Horodatage du moment où le problème a été observé pour la première fois et si le problème est récurrent

## Workflow de dépannage

Afin de dépanner APD, nous devons examiner 3 composants - vCenter, SCVM et hôte ESXi.

Ces étapes sont un workflow suggéré afin de localiser ou de réduire la source du symptôme Tous les chemins vers le bas observé. Veuillez noter que cette commande n'a pas besoin d'être suivie avec soin et que vous pouvez l'adapter en fonction des symptômes particuliers observés sur l'environnement du client.

## Archivage du serveur vCenter :

Se connecter à vCenter Server (VCS) et accéder à un **hôte affecté**

1. **Objets associés -> Machines virtuelles** et confirmer que StCtlVM est en cours d'exécution
2. **Objets associés -> Datastores** et confirmer si les datastores NFS montrent "inaccessible" . Si les datastores semblent **accessibles** et **accessibles**, vous pouvez essayer dans l'onglet **Résumé** de « Réinitialiser à Vert » l'événement APD et vérifier plus tard si l'alerte revient plus tard
3. **Surveiller -> Problèmes** et **Surveiller -> Événements** devraient fournir des renseignements sur le moment où la DPA a été détectée pour la première fois.

## Vérifie tous les StCtlVM :

Connectez-vous à **toutes** les **machines virtuelles StCtl** et vérifiez les pointeurs ci-dessous, vous pouvez utiliser [MobaXterm](#) le logiciel Cisco IOS.

1. Vérifiez si toutes les StCtlVM ont la même durée en utilisant **date** ou **ntpq -p**. Le décalage temporel sur StCtlVM peut entraîner des problèmes de synchronisation de la base de données de zookeeper et il est donc primordial de la synchroniser entre toutes les StCtlVM. L'information de signe astrick du serveur ntp indique que le NTP de votre SCVM est synchronisé.

```
root@SpringpathControllerPZTMTRSH7K:~# date
Tue May 28 12:47:27 PDT 2019
```

```
root@SpringpathControllerPZTMTRSH7K:~# ntpq -p -4
remote refid st t when poll reach delay offset jitter
=====
*abcdefghijkl.GNSS. 1 u 429 1024 377 225.813 -1.436 0.176
```

2. Si APD s'est produit au cours d'une **mise à niveau**, vous pouvez envisager de vérifier quels **StCtlVM** n'ont **pas été complètement mis à niveau** et en particulier d'identifier celui qui a échoué en dernier. Il est possible qu'il s'agisse de celui qui tenait le eth1:0 précédemment Utiliser **dpkg -l | grep -i springpath** pour identifier les StCtlVM non complètement mis à niveau car ils auront des paquets springpath de version mixte.

```
root@SpringpathControllerPZTMTRSH7K:~# dpkg -l | grep -i springpath
ii storfs-appliance 4.0.1a-33028 amd64 Springpath Appliance
ii storfs-asup 4.0.1a-33028 amd64 Springpath ASUP and SCH
ii storfs-core 4.0.1a-33028 amd64 Springpath Distributed Filesystem
ii storfs-fw 4.0.1a-33028 amd64 Springpath Appliance
ii storfs-mgmt 4.0.1a-33028 amd64 Springpath Management Software
ii storfs-mgmt-cli 4.0.1a-33028 amd64 Springpath Management Software
ii storfs-mgmt-hypervcli 4.0.1a-33028 amd64 Springpath Management Software
ii storfs-mgmt-ui 4.0.1a-33028 amd64 Springpath Management UI Module
ii storfs-mgmt-vcplugin 4.0.1a-33028 amd64 Springpath Management UI and vCenter Plugin
ii storfs-misc 4.0.1a-33028 amd64 Springpath Configuration
ii storfs-pam 4.0.1a-33028 amd64 Springpath PAM related modules
ii storfs-replication-services 4.0.1a-33028 amd64 Springpath Replication Services
ii storfs-restapi 4.0.1a-33028 amd64 Springpath REST Api's
```

```

ii storfs-robo 4.0.1a-33028 amd64 Springpath Appliance
ii storfs-support 4.0.1a-33028 amd64 Springpath Support
ii storfs-translations 4.0.1a-33028 amd64 Springpath Translations

```

3. Vérifiez si tous les services concernés sont *en cours d'exécution* `service_status.sh` : Certains des principaux services sont Springpath File System (storfs), SCVM Client (`scvmclient`), System Management Service (`stMgr`) ou Cluster IP Monitor (`cip-monitor`).

```

root@SpringpathController5L0GTCR8SA:~# service_status.sh
Springpath File System ... Running
SCVM Client ... Running
System Management Service ... Running
HyperFlex Connect Server ... Running
HyperFlex Platform Agnostic Service ... Running
HyperFlex HyperV Service ... Not Running
HyperFlex Connect WebSocket Server ... Running
Platform Service ... Running
Replication Services ... Running
Data Service ... Running
Cluster IP Monitor ... Running
Replication Cluster IP Monitor ... Running
Single Sign On Manager ... Running
Stats Cache Service ... Running
Stats Aggregator Service ... Running
Stats Listener Service ... Running
Cluster Manager Service ... Running
Self Encrypting Drives Service ... Not Running
Event Listener Service ... Running
HX Device Connector ... Running
Web Server ... Running
Reverse Proxy Server ... Running
Job Scheduler ... Running
DNS and Name Server Service ... Running
Stats Web Server ... Running

```

4. Si l'un de ces services ou d'autres services pertinents n'est pas actif, démarrez-le à l'aide de `start <serviceName>` par exemple : `start storfs` Vous pouvez consulter le script `service_status.sh` pour obtenir les noms de service . Faites une tête `-n25 /bin/service_status.sh` et identifiez le nom réel du service.

```

root@SpringpathController5L0GTCR8SA:~# head -n25 /bin/service_status.sh
#!/bin/bash
declare -a upstart_services=("Springpath File System:storfs"\
"SCVM Client:scvmclient"\
"System Management Service:stMgr"\
"HyperFlex Connect Server:hxmanager"\
"HyperFlex Platform Agnostic Service:hxSvcMgr"\
"HyperFlex HyperV Service:hxHyperVSvcMgr"\
"HyperFlex Connect WebSocket Server:zkupdates"\
"Platform Service:stNodeMgr"\
"Replication Services:replsvc"\
"Data Service:stDataSvcMgr"\
"Cluster IP Monitor:cip-monitor"\
"Replication Cluster IP Monitor:repl-cip-monitor"\
"Single Sign On Manager:stSSOMgr"\
"Stats Cache Service:carbon-cache"\
"Stats Aggregator Service:carbon-aggregator"\
"Stats Listener Service:statsd"\
"Cluster Manager Service:exhibitor"\
"Self Encrypting Drives Service:sedsvc"\
"Event Listener Service:storfsevents"\
"HX Device Connector:hx_device_connector");
declare -a other_services=("Web Server:tomcat8"\
"Reverse Proxy Server:nginx"\
"Job Scheduler:cron"\

```

```
"DNS and Name Server Service:resolvconf");
```

5. Identifiez quelle StCtlVM contient l'adresse IP du cluster de stockage (eth1:0) à l'aide de **ifconfig -a** Si aucun StCtlVM ne contient cette adresse IP, les storfs ne sont peut-être pas exécutés sur un ou plusieurs noeuds.

```
root@help:~# ifconfig
eth0:mgmtip Link encap:Ethernet HWaddr 00:50:56:8b:4c:90
inet addr:10.197.252.83 Bcast:10.197.252.95 Mask:255.255.255.224
UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
```

6. Vérifiez si StCtlVM est en contact avec CRMMaster et si zookeeper service est actif et en cours d'exécution **echo svr | nc localhost 2181** et vérifiez si le mode est Leader, Follower ou Autalone et si les connexions > 0

```
root@help:~# echo svr | nc localhost 2181
Zookeeper version: 3.4.12-d708c3f034468a4da767791110332281e04cf6af, built on 11/19/2018
21:16 GMT
Latency min/avg/max: 0/0/137
Received: 229740587
Sent: 229758548
Connections: 13
Outstanding: 0
Zxid: 0x140000526c
Mode: leader
Node count: 3577
```

**statut d'exposant de service ou ps -ef | exposant grep -i**

```
root@help:~# service exhibitor status
exhibitor start/running, process 12519
root@help:~# ps -ef | grep -i exhibitor
root 9765 9458 0 13:19 pts/14 00:00:00 grep --color=auto -i exhibitor
root 12519 1 0 May19 ? 00:05:49 exhibitor
```

En cas d'erreur ou de service non en cours d'exécution, vous pouvez vérifier les journaux ci-dessous et essayer de démarrer le service zookeeper **/var/log/springpath/exhibitor.log** et **/var/log/springpath/stMgr.log** **service expositor commence** à démarrer zookeeper service

7. Vérifier si VC est accessible à partir de toutes les machines virtuelles StCtl informations de cluster **stcli | grep -i « url »** pour afficher l'URL utilisée contenant soit FQDN soit IP de VC. Vérifiez la connectivité au circuit virtuel à l'aide de la commande **ping <VC>**

```
root@help:~# stcli cluster info | grep -i "url"
vCenterUrl: https://10.197.252.101
vCenterURL: 10.197.252.101
root@help:~# ping 10.197.252.101
PING 10.197.252.101 (10.197.252.101) 56(84) bytes of data.
64 bytes from 10.197.252.101: icmp_seq=1 ttl=64 time=0.435 ms
```

8. Confirmer si DNS est accessible en cas de cluster utilisant FQDN **les dns des services stcli indiquent** la liste des serveurs DNS configurés sur StCtlVM. Testez la connectivité et la résolution aux serveurs DNS à l'aide de la commande **ping <DNS\_IP>** et **host <FQDN> <DNS\_IP>**

```
root@help:~# stcli services dns show
1.1.128.140
root@help:~# ping 1.1.128.140
PING 1.1.128.140 (1.1.128.140) 56(84) bytes of data.
64 bytes from 1.1.128.140: icmp_seq=1 ttl=244 time=1.82 ms
```

9. Confirmer si toutes les StCtlVM ont la même quantité d'entrées iptables : **iptables -L | wc -l**. En cas de non-concordance, veuillez ouvrir un dossier TAC.

```
root@SpringpathControllerI51U7U6QZX:~# iptables -L | wc -l
48
```

10. Quel est l'état et l'état actuels du cluster ? **informations de cluster stcli | moins** ou **informations de cluster stcli | grep -i « actif/état/indisponible »** si vous essayez de trouver les noeuds spécifiques qui apparaissent avec le stockage indisponible.

```

root@SpringpathControllerI51U7U6QZX:~# stcli cluster info | grep -i
"active\|state\|unavailable"
locale: English (United States)
state: online
upgradeState: ok
healthState: healthy
state: online
state: 1
activeNodes: 3
state: online

```

**stcli cluster storage-summary —détail** Pour obtenir les détails du cluster de stockage

```

root@SpringpathControllerI51U7U6QZX:~# stcli cluster storage-summary --detail
address: 10.197.252.106
name: HX-Demo
state: online
uptime: 185 days 12 hours 48 minutes 42 seconds
activeNodes: 3 of 3
compressionSavings: 85.45%
deduplicationSavings: 0.0%
freeCapacity: 4.9T
healingInfo:
inProgress: False
resiliencyDetails:
current ensemble size:3
# of caching failures before cluster shuts down:3
minimum cache copies remaining:3
minimum data copies available for some user data:3
minimum metadata copies available for cluster metadata:3
# of unavailable nodes:0
# of nodes failure tolerable for cluster to be available:1
health state reason:storage cluster is healthy.
# of node failures before cluster shuts down:3
# of node failures before cluster goes into readonly:3
# of persistent devices failures tolerable for cluster to be available:2
# of node failures before cluster goes to enospace warn trying to move the existing
data:na
# of persistent devices failures before cluster shuts down:3
# of persistent devices failures before cluster goes into readonly:3
# of caching failures before cluster goes into readonly:na
# of caching devices failures tolerable for cluster to be available:2
resiliencyInfo:
messages:
Storage cluster is healthy.
state: 1
nodeFailuresTolerable: 1
cachingDeviceFailuresTolerable: 2
persistentDeviceFailuresTolerable: 2
zoneResInfoList: None
spaceStatus: normal
totalCapacity: 5.0T
totalSavings: 85.45%
usedCapacity: 85.3G
zkHealth: online
clusterAccessPolicy: lenient
dataReplicationCompliance: compliant
dataReplicationFactor: 3

```

## 11. Quels datastores sont montés et disponibles ?

```

root@bsv-hxaf220m5-sc-4-3:~# stcli datastore list
-----
virtDatastore:
  status:
    EntityRef(idtype=None, confignum=None, type=6, id='235ea35f-6c85-9448-bec7-
06f03b5adf16', name='bsv-hxaf220m5-hv-4-3.cisco.com'):

```

```

    accessible: True
    mounted: True
    EntityRef(idtype=None, confignum=None, type=6, id='d124203c-3d9a-ba40-a229-4dffbe96ae13', name='bsv-hxaf220m5-hv-4-2.cisco.com'):
    accessible: True
    mounted: True
    EntityRef(idtype=None, confignum=None, type=6, id='e85f1980-b3c7-a440-9f1e-20d7a1110ae6', name='bsv-hxaf220m5-hv-4-1.cisco.com'):
    accessible: True
    mounted: True

```

12. Au cas où les commandes stcli prendraient trop de temps ou échoueraient, vous pouvez essayer les commandes sysmtool suivantes (Ne pas utiliser si stcli fonctionne) **sysmtool**
  - ns cluster —cmd info sysmtool —ns cluster —cmd health detail sysmtool —ns datastore —cmd list

## StCtlVM : StCtlVM d'un hôte ESXi affecté

Se connecter à la machine virtuelle StCtl de l'hôte ESXi concerné

1. Vérifiez la connectivité à l'IP du cluster de stockage (eth1:0) et aux autres serveurs du réseau de stockage (eth1 sur StCtlVM)
 

Exécuter les *informations de cluster stcli | grep -i -B 1 « stctl\hyperviseur »* pour identifier toutes les adresses IP de gestion ESXi, StCtlVM eth0 (Mgmt) et StCtlVM eth1 (données de stockage) participant respectivement au cluster. Testez la connectivité *ping -I eth1 [-M do -s 8972] <adresse IP cible>*, le test des trames Jumbo entre ESXi VMK1 et SCVM eth1.
2. Si le problème n'est toujours pas identifié, vous pouvez consulter les journaux suivants */var/log/springpath/debug-storfs.log* Vérifiez si des paniques, des erreurs de SEG ou des événements critiques *grep -ai « segmentation\|critique\|panic » debug-storfs.log/var/log/springpath/stmgr.log* vérifiez si un problème de mémoire insuffisante présente *grep -i « oom\|out of mem » /var/log/kern.log*
3. En fin de compte, vous pouvez essayer de redémarrer la machine virtuelle StCtl du noeud qui connaît toujours le problème et vérifier si le problème persiste.

## Vérifie dans l'hôte ESXi :

Connectez-vous à un hôte ESXi affecté via SSH et effectuez les opérations suivantes :

1. **esxcli storage nfs list** ou **Esxcfg-nas -l** POUR répertoirer les data stores NFS actuellement montés et s'ils sont accessibles
 

```

[root@bsv-hx220m5-hv-4-3:~] esxcli storage nfs list
Volume Name  Host                               Share                               Accessible
Mounted  Read-Only  isPE  Hardware Acceleration
-----  -
test      8352040391320713352-8294044827248719091  192.168.4.1:test                  true
true     false     false Supported
sradzevi   8352040391320713352-8294044827248719091  192.168.4.1:sradzevi             true
true     false     false Supported
[root@bsv-hx220m5-hv-4-3:~] esxcfg-nas -l
test is 192.168.4.1:test from 8352040391320713352-8294044827248719091 mounted available
sradzevi is 192.168.4.1:sradzevi from 8352040391320713352-8294044827248719091 mounted available

```

Vous pouvez également confirmer à partir de */etc/vmware/esx.conf* pour vérifier la cohérence de la configuration ESXi sur les data stores montés sur NFS, à l'aide de la commande **cat**

*/etc/vmware/esx.conf | grep -I nas*

2. Vérifiez */var/log/vmkernel.log* et recherchez par exemple un état d'échec, des problèmes de montage ou une erreur autour de l'horodatage identifié lors des étapes précédentes
3. Vérifier l'état du proxy IOVisor/NFS/SCVMClient Vérifiez si le **service** est en cours d'exécution sur ESXi à l'aide de la commande */etc/init.d/scvmclient status [Facultatif]* Vous pouvez vérifier si des connexions ouvertes utilisent **la liste de connexions ip réseau esxcli | grep -i « proto/scvmclient »** Confirmez si SCVMClient VIB est la même version que votre cluster HX, **liste des vidéos du logiciel Esxcli | grep -i printemps**

```
[root@bsv-hx220m5-hv-4-3:~] esxcli software vib list | grep -i spring
scvmclient                3.5.1a-31118                Springpath
VMwareAccepted           2018-12-13
stHypervisorSvc          3.5.1a-31118                Springpath
VMwareAccepted           2018-12-06
vmware-esx-STFSNasPlugin 1.0.1-21                    Springpath
VMwareAccepted           2018-11-16
```

Consultez le fichier */var/log/scvmclient.log* pour savoir s'il y a des erreurs, à savoir « impossible d'obtenir la carte de cluster » Vous pouvez redémarrer le service SCVMClient si nécessaire via *etc/init.d/scvmclient restart*

4. Vérifier la connectivité réseau avec les autres hôtes ESXi sur le réseau vmk1, en particulier vers le cluster de stockage IP eth1:0*esxcfg-vmknic -I* pour obtenir des informations sur les détails de la carte réseau vmk, par exemple IP, masque et MTU *vmkping -I vmk1 [-v -s 8972] -d <adresse IP cible>* à test de la connectivité [éventuellement avec des trames jumbo] entre les hôtes ESXi sur le réseau de données du contrôleur
5. **la plate-forme matérielle esxcli** obtient le numéro de série du serveur utilisé sur le nom de StCtlVm et peut vous aider à identifier rapidement sur quel hôte une StCtlVM spécifique est exécutée.