

# Resolución de problemas de trayectos hacia abajo (APD) de almacenamiento Hyperflex NFS

## Contenido

[Introducción](#)

[¿Cómo se montan los almacenes de datos HX en ESXi?](#)

[Todos los trayectos hacia abajo](#)

[Descripción de problemas](#)

[Resolución de problemas de flujo de trabajo](#)

[Protege el servidor vCenter:](#)

[Protege todas las StCtIVM:](#)

[StCtIVM: StCtIVM de un host ESXi afectado](#)

[Verifica en el host ESXi:](#)

## Introducción

Este documento le proporciona una comprensión rápida y pasos de troubleshooting que se pueden realizar para evaluar el origen del problema si ve el mensaje de error "NFS all paths down" en vCenter con el cual se integra el clúster Hyperflex.

## ¿Cómo se montan los almacenes de datos HX en ESXi?

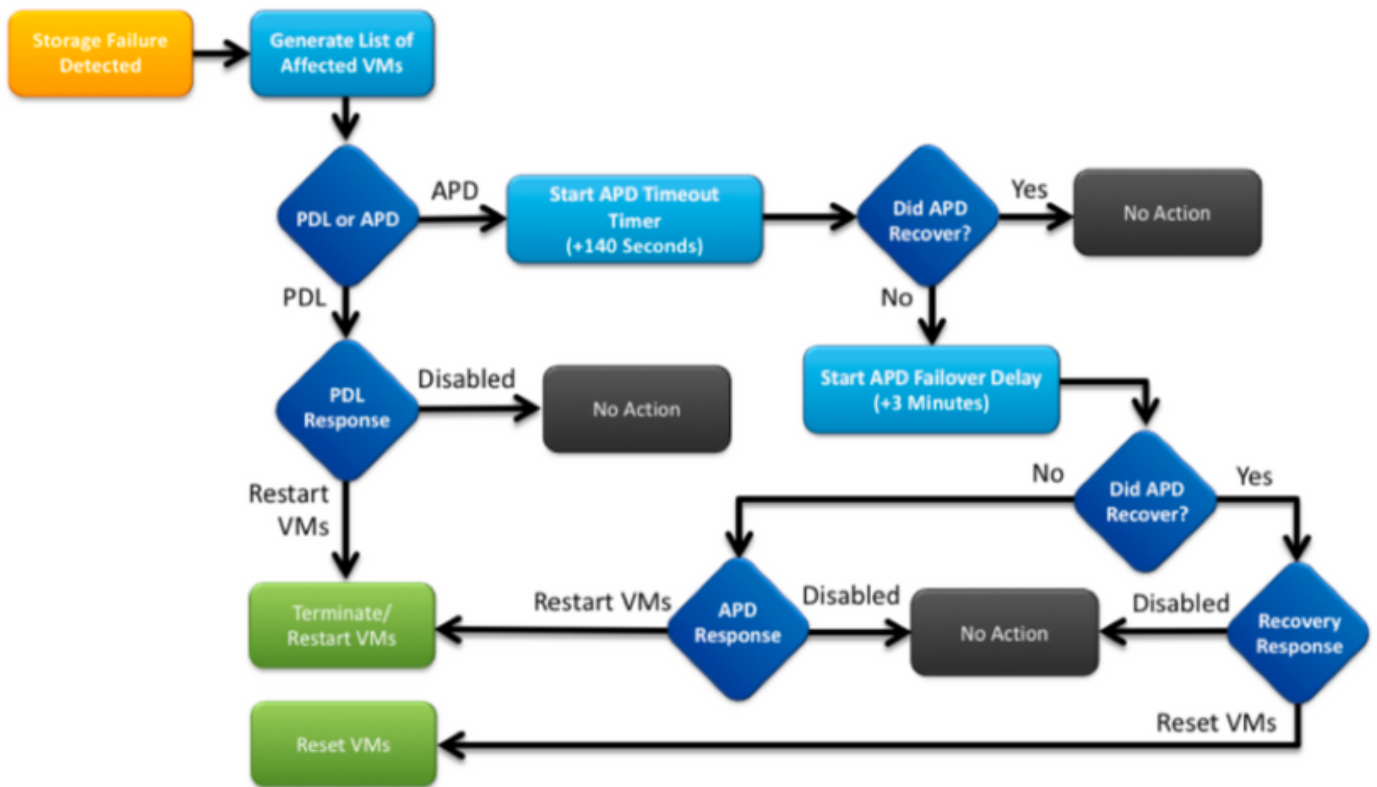
- Los almacenes de datos Hyperflex se montan en los hosts ESXi como montajes NFS, para montar un almacén de datos NFS necesitamos la IP del servidor NFS que en nuestro caso es la interfaz virtual flotante eth1:0.
- El clúster Hyperflex aprovecha el uso de **IP flotante** virtual tanto para la administración (eth0:mgmtip) como para los **datos de almacenamiento (eth1:0) en los que cada IP se asignará a una VM del controlador de almacenamiento determinada (StCtIVM)**. Tenga en cuenta que pueden terminar en StCtIVM diferentes.
- La importancia de esto es que la **IP de datos de almacenamiento de clúster (eth1:0) es la que se utiliza para montar el almacén de datos** creado en el **hyperflex** cluster. Por lo tanto, es esencial que se le **asigne y se pueda alcanzar** desde todos los **nodos** del clúster.
- Tenga en cuenta que en caso de que se produzca un error en StCtIVM que actualmente posee eth1:0 virtual IP, debería "migrar" a otra StCtIVM disponible que funcione de forma similar a un FHRP (First Hop Redundancy Protocol).

## Todos los trayectos hacia abajo

- APD significa que el host no puede alcanzar el almacenamiento y que no hay ningún código SCSI Perdido de dispositivo permanente (PDL) devuelto desde la matriz de almacenamiento.
- Como no sabe si la pérdida es temporal o no, seguirá intentando establecer la comunicación para más 140 de forma predeterminada (tiempo de espera) + 3min (Retraso para la conmutación por fallas) el host ESXi comienza a fallar cualquier tráfico de E/S que no sea de

máquina virtual que se envía al dispositivo de almacenamiento.

- Para obtener más detalles con respecto a APD en el entorno vSphere 6.x, consulte el artículo de VMware KB [Pérdida permanente de dispositivos \(PDL\) y All-Paths-Down \(APD\) en vSphere 5.x y 6.x](#)
- La siguiente imagen explica cada etapa intermitente:



Un mensaje de error típico en vCenter será el siguiente.

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

## Descripción de problemas

Una vez que vea las alertas APD en sus hosts, obtenga la siguiente información para entender mejor la descripción del problema:

- Si uno/varios/todos los hosts se han visto afectados, y si algunos de los hosts en particular se han visto afectados
- Si se realizaron cambios anteriormente (configuración/actualización/etc.)
- Marca de tiempo sobre cuándo se observó el problema por primera vez y si el problema es recurrente

## Resolución de problemas de flujo de trabajo

Para solucionar problemas de APD, necesitamos analizar 3 componentes: vCenter, SCVM y host ESXi.

Estos pasos son un flujo de trabajo sugerido para identificar o restringir el origen del síntoma de All Paths Down observado. Tenga en cuenta que este pedido no tiene que ser seguido meticulosamente y puede ser adecuado según los síntomas particulares observados en el entorno del cliente.

## Protege el servidor vCenter:

Conéctese a vCenter Server (VCS) y navegue hasta un **host afectado**

1. **Objetos relacionados -> Máquinas virtuales** y confirme que StCtlVM está en funcionamiento
2. **Objetos relacionados -> Almacenes de datos** y confirmar si los almacenes de datos NFS muestran "inaccesible". Si los almacenes de datos parecen estar **accesibles** y el estado, puede probar en la pestaña **Resumen** para "Restablecer a verde" el evento APD y posteriormente verificar si la alerta vuelve a aparecer más tarde
3. **Monitor -> Problemas y Monitor -> Eventos deben proporcionar información sobre cuándo se detectó por primera vez la APD.**

## Protege todas las StCtlVM:

Conéctese a **todas** las VM StCtl y verifique los punteros siguientes, puede utilizar [MobaXterm](#) software.

1. Verifique si todas las StCtlVMs tienen el mismo tiempo usando **date** o **ntpq -p**. La desviación de tiempo en StCtlVM puede provocar problemas con la sincronización de la base de datos de zookeeper y, por lo tanto, es primordial tenerla sincronizada entre todas las StCtlVM. El intento de firma del servidor ntp indica que se sincroniza el NTP de su SCVM.

```
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 se produjo un APD durante una **actualización**, podría considerar verificar qué **StCtlVM no se han actualizado completamente** e identificar en particular el que falló por última vez. Es posible que fuera el que sostenía el eth1:0 anteriormente Usar **dpkg -l | grep -i springpath** para identificar las StCtlVM que no se han actualizado completamente, ya que tendrán paquetes de tramas de versiones mixtas.

```
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. Verifique si todos los servicios relevantes están **ejecutando service\_status.sh**: Algunos de los servicios principales son el Sistema de archivos Springpath (storfs), Cliente SCVM (*scvmclient*), Servicio de administración del sistema (*stMgr*) o Monitor IP de clúster (*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 alguno de estos u otros servicios relevantes no está activo, comience a utilizar **start <serviceName>** eg: **start storfs** Puede consultar el script `service_status.sh` para obtener los nombres de servicio . Haga un `head -n25 /bin/service_status.sh` e identifique el nombre real del servicio.

```
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. Identifique **qué StCtlVM** contiene el **cluster de almacenamiento IP (eth1:0)** usando **ifconfig -a**. Si no hay StCtlVM que contenga esa IP, es posible que los almacenes no se estén ejecutando en uno o más nodos.

```
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. Verifique si StCtlVM está en contacto con CRMMaster y si **zookeeper service** está activo y en ejecución **echo svr | nc localhost 2181** y verifique si el modo es líder, seguidor o independiente y si las conexiones **> 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
```

**estado del expositor del servicio o ps -ef | expositor 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 caso de que se produzca algún error o servicio que no se esté ejecutando, puede verificar los siguientes registros e intentar iniciar el servicio zookeeper

**/var/log/springpath/exhibitor.log** y **/var/log/springpath/stMgr.log**el **expositor de servicio** comienza a iniciar el servicio zookeeper

7. Verifique si **VC alcanzable** desde todas las **VMs StCtl** información de clúster de STcli | **grep -i "url"** para mostrar la URL utilizada que contiene FQDN o IP de VC. Verifique la conectividad con el VC usando **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. Confirme si **DNS** es **accesible** en caso de que el clúster utilice **FQDN stcli services dns show** para mostrar los servidores configurados de DNS en StCtlVM. Pruebe la **conectividad** y la **resolución** a los servidores **DNS** mediante **ping <DNS\_IP>** y **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. Confirme si todas **StCtlVM** tienen la misma cantidad de **entradas iptables: iptables -L | wc -l**. **En caso de que no coincidan, abra un caso del TAC.**

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

10. ¿Cuáles son el estado y el estado actuales del clúster? **información de clúster de STcli | less** o **stcli cluster info | grep -i "active|state|no disponible"** si intenta encontrar qué nodos concretos aparecen con almacenamiento no disponible.

```

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 --detail** para obtener los detalles del clúster de almacenamiento

```

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. Qué almacenes de datos están montados y 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. En caso de que los comandos stcli tarden demasiado o fallen, puede probar los siguientes comandos sysmtool (No lo use si stcli funciona) **sysmtool —ns cluster —cmd info sysmtool —ns cluster —cmd health-detail sysmtool —ns datastore —cmd list**

## StCtIVM: StCtIVM de un host ESXi afectado

Conectar con la StCtIVM del host ESXi afectado

1. Verifique la conectividad con la IP del clúster de almacenamiento (eth1:0) y con otros servidores en la red de almacenamiento (eth1 en StCtIVM)  
Ejecutar **información de clúster de stcli | grep -i -B 1 "stctl\|hypervisor"** para identificar todas las IP de administración de ESXi, StCtIVM eth0 (Mgmt) y StCtIVM eth1 (datos de almacenamiento) que participan respectivamente en el clúster. Pruebe el **ping de conectividad -I eth1 [-M do -s 8972] <dirección IP de destino>, prueba de tramas Jumbo entre ESXI VMK1 y SCVM eth1.**
2. Si el problema aún no se ha señalado, es posible que tenga que examinar los siguientes registros **/var/log/springpath/debug-storfs.log** Verifique si hay algún pánico, falla de s o eventos críticos **grep -ai "segmentation\|crítico\|pánico" debug-storfs.log/var/log/springpath/stmgr.log** Verifique si el problema de memoria fuera de la memoria presenta **grep -i "oom\|out of mem" /var/log/kern.log**
3. En última instancia, puede intentar reiniciar la StCtIVM del nodo que aún experimenta el problema y verificar si el problema persiste.

## Verifica en el host ESXi:

Conéctese a un host ESXi afectado a través de SSH y realice las siguientes acciones:

1. **esxcli storage nfs list oEsxcfg-nas -l** para enumerar los almacenes de datos NFS montados actualmente y si son accesibles

```

[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

```

También puede confirmar desde **/etc/vmware/esx.conf** para verificar la consistencia en la

configuración de ESXi en los datastores montados de NFS, usando el comando **cat /etc/vmware/esx.conf | grep -I nas**

2. Verifique **/var/log/vmkernel.log** y busque por ejemplo el estado fallido, los problemas de montaje o el error en la marca de tiempo identificada en los pasos anteriores
3. Verifique el estado de IOvisor/NFS Proxy/SCVMClient Compruebe si el servicio se está ejecutando en ESXi mediante el comando **etc/init.d/scvmclient status [Opcional]** Puede verificar si hay conexiones abiertas usando la **lista de conexiones ip de red esxcli | grep -i "proto|scvmclient"** Confirme si SCVMClient VIB es la misma versión que su clúster HX, **lista de vib del software Esxcli | grep -i primavera**

```
[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
```

Verifique **/var/log/scvmclient.log** para ver si hay algún error, a saber, "no se puede obtener clustermap" Puede reiniciar el servicio SCVMClient si es necesario a través de **etc/init.d/scvmclient restart**

4. Verifique la conectividad de red con otros hosts ESXi en la red vmk1, particularmente con el clúster de almacenamiento IP **eth1:0esxcfg-vmknic -I** para obtener información sobre los detalles de vmk nic, por ejemplo, IP, máscara y MTU **vmkping -I vmk1 [-v -s 8972] -d <dirección IP de destino>** aprobar conectividad [opcionalmente con tramas jumbo] entre hosts ESXi en la red de datos del controlador
5. **la plataforma de hardware esxcli obtiene** el servidor SN que se utiliza en el nombre de StCtlVM y puede ayudarle a identificar rápidamente en qué host se ejecuta una StCtlVM específica.