

# Troubleshooting de NFS de Armazenamento Hiperflex - problema com todos os caminhos para baixo (APD)

## Contents

[Introduction](#)

[Como os datastores HX são montados no ESXI?](#)

[Todos os caminhos para baixo](#)

[Descrição do problema](#)

[Troubleshooting do Fluxo de Trabalho](#)

[Verifica no vCenter Server:](#)

[Verifica todas as VMs StCtl:](#)

[StCtlVM: StCtlVM de um host ESXi afetado](#)

[Verifica no host ESXi:](#)

## Introduction

Este documento fornece compreensão rápida e etapas de solução de problemas que podem ser executadas para avaliar a origem do problema se você vir a mensagem de erro "NFS all paths down" no vCenter ao qual o cluster do Hyperflex está integrado.

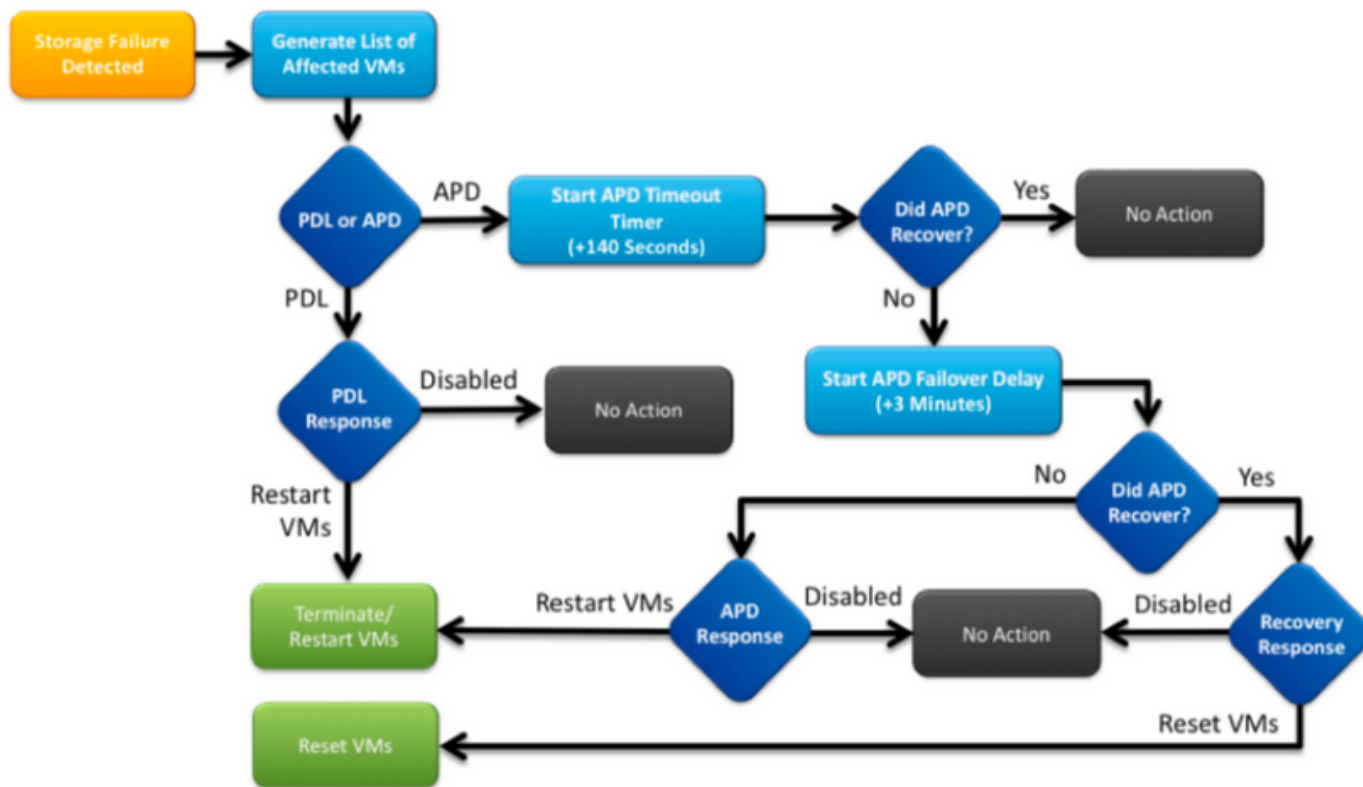
## Como os datastores HX são montados no ESXI?

- Os Hyperflex Datastores são montados nos hosts ESXi como montagens NFS, para montar um armazenamento de dados NFS, precisamos do IP do NFS Server que, no nosso caso, é a interface virtual flutuante eth1:0.
- O cluster hiperflex aproveita o uso de **IP virtual flutuante** tanto para gerenciamento (eth0:mgmtip) quanto **dados** de armazenamento (**eth1:0**) nos quais cada IP será atribuído a uma VM de Controlador de Armazenamento específica (**StCtlVM**). Observe que eles podem acabar em StCtlVMs diferentes.
- A importância disso é que o **IP de dados de armazenamento de cluster (eth1:0) é o usado para montar o(s) datastore(s) criado(s) no cluster do Hyperflex**. Assim, é essencial que ele seja **atribuído** e **acessível** de todos os **nós** do cluster.
- Observe que, em caso de falha da StCtlVM que atualmente possui IP virtual eth1:0, ela deve "migrar" para outra StCtlVM disponível que funcione de forma semelhante a um FHRP (First Hop Redundancy Protocol).

## Todos os caminhos para baixo

- APD significa que o host não pode acessar o armazenamento e que não há código SCSI PDL (Permanent Device Lost, dispositivo permanente perdido) retornado da matriz de armazenamento.

- Como não sabe se a perda é temporária ou não, ele continuará tentando estabelecer comunicação por mais de 140 s por padrão (tempo limite) + 3 minutos (atraso para failover), o ESXi Host começa a falhar em qualquer tráfego de I/O de máquina não virtual que está sendo enviado ao dispositivo de armazenamento.
- Para obter mais detalhes sobre APD no ambiente vSphere 6.x, consulte o artigo do KB do VMware [Perda de dispositivo permanente \(PDL\) e All-Paths-Down \(APD\) no vSphere 5.x e 6.x](#)
- A imagem abaixo explica cada estágio intermitente:



Uma mensagem de erro típica no vCenter será a seguinte.

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

## Descrição do problema

Depois de ver os alertas APD sobre seus hosts, obtenha as informações abaixo para entender melhor a descrição do problema:

- Se um/vários/todos os hosts tiverem impacto, e se alguns dos hosts específicos tiverem impacto
- Se alguma alteração tiver sido feita anteriormente (configuração/atualização/etc.)
- O carimbo de data e hora em que o problema foi observado pela primeira vez e se o problema é recorrente

## Troubleshooting do Fluxo de Trabalho

Para solucionar problemas de APD, precisamos examinar 3 componentes - vCenter, SCVM e host ESXi.

Essas etapas são um fluxo de trabalho sugerido para apontar ou restringir a origem do sintoma All Paths Down observado. Observe que este pedido não precisa ser seguido meticulosamente e você pode adequá-lo de acordo com os sintomas específicos observados no ambiente do cliente.

## Verifica no vCenter Server:

Conecte-se ao vCenter Server (VCS) e navegue até um **host afetado**

1. **Objetos relacionados -> Máquinas virtuais** e confirme se o StCtlVM está ativo e em execução
2. **Objetos relacionados -> Datastores** e confirme se os datastores NFS mostram "inacessível". Se os armazenamentos de dados parecerem **acessíveis** e o status, você pode tentar na guia **Resumo** para "Redefinir para verde" o evento APD e verificar mais tarde se o alerta aparece novamente mais tarde
3. **Monitor -> Issues and Monitor -> Events** deve fornecer informações sobre quando o APD foi detectado pela primeira vez.

## Verifica todas as VMs StCtl:

Conecte-se a **todas** as VMs StCtl e verifique os indicadores abaixo. Você pode usar [MobaXterm](#) software.

1. Verifique se todas as StCtlVMs têm o mesmo tempo usando **date** ou **ntpq -p**. O desvio de tempo em StCtlVM pode levar a problemas com a sincronização do banco de dados de zookeeper e, portanto, é fundamental tê-lo em sincronização entre todas as StCtlVMs. O sinal de entrada do servidor ntp indica que o NTP de seu SCVM está sincronizado.

```
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. Se o APD ocorreu durante uma **atualização**, você pode considerar verificar quais VMs de **StCtl não foram completamente atualizadas** e, particularmente, identificar a que falhou pela última vez. É possível que tenha sido o que segurava os dentes 1:0 anteriormente Usar **dpkg -l | grep -i springpath** para identificar as StCtlVMs que não foram completamente atualizadas, pois elas terão pacotes de springpath de versão mista.

```
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 se todos os serviços relevantes estão **executando service\_status.sh**: Alguns dos principais serviços são 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. Se algum desses ou de outros serviços relevantes não estiver ativo, inicie-o **usando start <serviceName>** eg: **start storfs** Você pode consultar o script `service_status.sh` para obter os nomes dos serviços . Faça um `head -n25 /bin/service_status.sh` e identifique o nome real do serviço.

```

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 qual StCtlVM contém o cluster de armazenamento IP (eth1:0) usando *ifconfig -a*. Se nenhuma StCtlVM contiver esse IP, possivelmente os storfs não estarão em execução em um ou mais nós.

```
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 se StCtlVM está em contato com o CRMMaster e se o serviço zookeeper está ativo e em execução *echo svr | nc localhost 2181* e verifique se mode é Leader, Follower ou Standalone e if connections > 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
```

#### status do exibidor de serviço ou 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
```

Em caso de erro ou serviço que não esteja em execução, você pode verificar os registros abaixo e tentar iniciar o serviço de manutenção */var/log/springpath/exhibitor.log*

e */var/log/springpath/stMgr.log* Expositor de serviço começa a iniciar o serviço de zootecnista

7. Verifique se o VC pode ser acessado de todas as VMs StCtl informação de cluster stcli | *grep -i "url"* para mostrar a URL usada contendo FQDN ou IP do VC. Verifique a conectividade com o VC usando o *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 se DNS está acessível no caso do cluster que usa FQDN *stcli services dns mostram* para listar os servidores DNS configurados em StCtlVM. Teste a conectividade e resolução para servidores DNS usando *ping <DNS\_IP>* e *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 se todas as VMs StCtl têm a mesma quantidade de entradas iptables: *iptables -L | wc -l*. Em caso de incompatibilidade, abra um caso no TAC.

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

10. Qual é o status e a integridade do cluster atual? *informação de cluster stcli | menos* ou *informações do cluster stcli | grep -i "ative|state|unavailable"* se tentar encontrar os nós

específicos que aparecem com armazenamento indisponível.

```
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 —detalhe** obter os detalhes do cluster de armazenamento

```
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. Quais datastores estão montados e disponíveis

```
root@bsv-hxaf220m5-sc-4-3:~# stcli datastore list
-----
virtDatastore:
  status:
    EntityRef(idtype=None, confignum=None, type=6, id='235ea35f-6c85-9448-bec7-
```

```

06f03b5adfl16', 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. Caso os comandos stcli demorem muito ou falhem, você pode tentar os seguintes comandos sysmtool(Não usar se o stcli funcionar) **sysmtool —ns cluster —cmd info sysmtool —ns cluster —cmd health detail sysmtool —ns datastore —cmd list**

## StCtlVM: StCtlVM de um host ESXi afetado

Ligar à StCtlVM do anfitrião ESXi afetado

1. Verifique a conectividade com o IP do cluster de armazenamento (eth1:0) e com outros servidores na rede de armazenamento (eth1 em VMsCtl)  
Executar **informações do cluster stcli | grep -i -B 1 "stctl|hypervisor"** para identificar todos os IP de gerenciamento ESXi, StCtlVM eth0 (Mgmt) e StCtlVM eth1 (dados de armazenamento) que participam respectivamente do cluster  
Teste a conectividade **ping -I eth1 [-M do -s 8972] <target IP address>, Jumbo frames test entre ESXI VMK1 e SCVM eth1.**
2. Se o problema ainda não tiver sido identificado, você pode consultar os seguintes registros **/var/log/springpath/debug-storfs.log** Verifique se há pânico, falha de seg ou eventos críticos **grep -ai "segmentation|Critical|Pânico" debug-storfs.log/var/log/springpath/stmgr.log** Verifique se o problema de memória não está presente **grep -i "oom|out of mem" /var/log/kern.log**
3. Em última análise, você pode tentar reinicializar a StCtlVM do nó ainda com o problema e verificar se o problema persiste.

## Verifica no host ESXi:

Conecte-se a um host ESXi afetado via SSH e execute as seguintes ações:

1. **lista nfs de armazenamento esxcli ouEsxcfg-nas -l** para listar os datastores NFS atualmente montados e se eles estão acessíveis

```

[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

```

Você também pode confirmar em **/etc/vmware/esx.conf** para verificar a consistência na configuração do ESXi nos datastores montados em NFS, usando o comando **cat**

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

2. Verifique */var/log/vmkernel.log* e procure por exemplo o estado de falha, problemas de montagem ou erro em torno do timestamp identificado nas etapas anteriores
3. Verifique o status de IOVisor/NFS Proxy/SCVMClient Verifique se o **serviço** está em execução no ESXi usando o comando */etc/init.d/scvmclient status [Opcional]* Você pode verificar se há conexões abertas usando a *lista de conexões ip de rede esxcli | grep -i "proto|scvmclient"* Confirme se o SCVMClient VIB é da mesma versão do cluster HX, *lista de vib do 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 se existem erros, nomeadamente "impossível obter o mapa de clusters" Você pode reiniciar o serviço SCVMClient, se necessário, por meio de *etc/init.d/scvmclient restart*

4. Verificar a conectividade de rede com outros hosts ESXi na rede vmk1, particularmente com o cluster de armazenamento  
IP *eth1:0esxcfg-vmknic -I* para obter informações sobre os detalhes da placa de rede vmk, como IP, máscara e MTU *vmkping -I vmk1 [-v -s 8972] -d <endereço IP de destino>* para Testar a conectividade [opcionalmente com quadros jumbo] entre hosts ESXi na rede de dados do controlador
5. **a plataforma de hardware esxcli obtém** a SN do servidor usada no nome do StCtIVm e pode ajudá-lo a identificar rapidamente em que host um StCtIVM específico está em execução.