

Hyperflex Storage NFS all paths down(APD)の問題のトラブルシューティング

内容

[概要](#)

[HXデータストアはESXIにどのようにマウントされますか。](#)

[すべてのパスダウン](#)

[問題の説明](#)

[トラブルシューティングワークフロー](#)

[vCenter Serverのチェック：](#)

[すべてのStCtIVMをチェックインします。](#)

[StCtIVM:該当するESXiホストのStCtIVM](#)

[ESXiホストを確認します。](#)

概要

このドキュメントでは、Hyperflexクラスタが統合されているvCenterで「NFS all paths down」エラーメッセージが表示された場合に、問題の原因を評価するために実行できる手順を簡単に説明し、トラブルシューティングを行います。

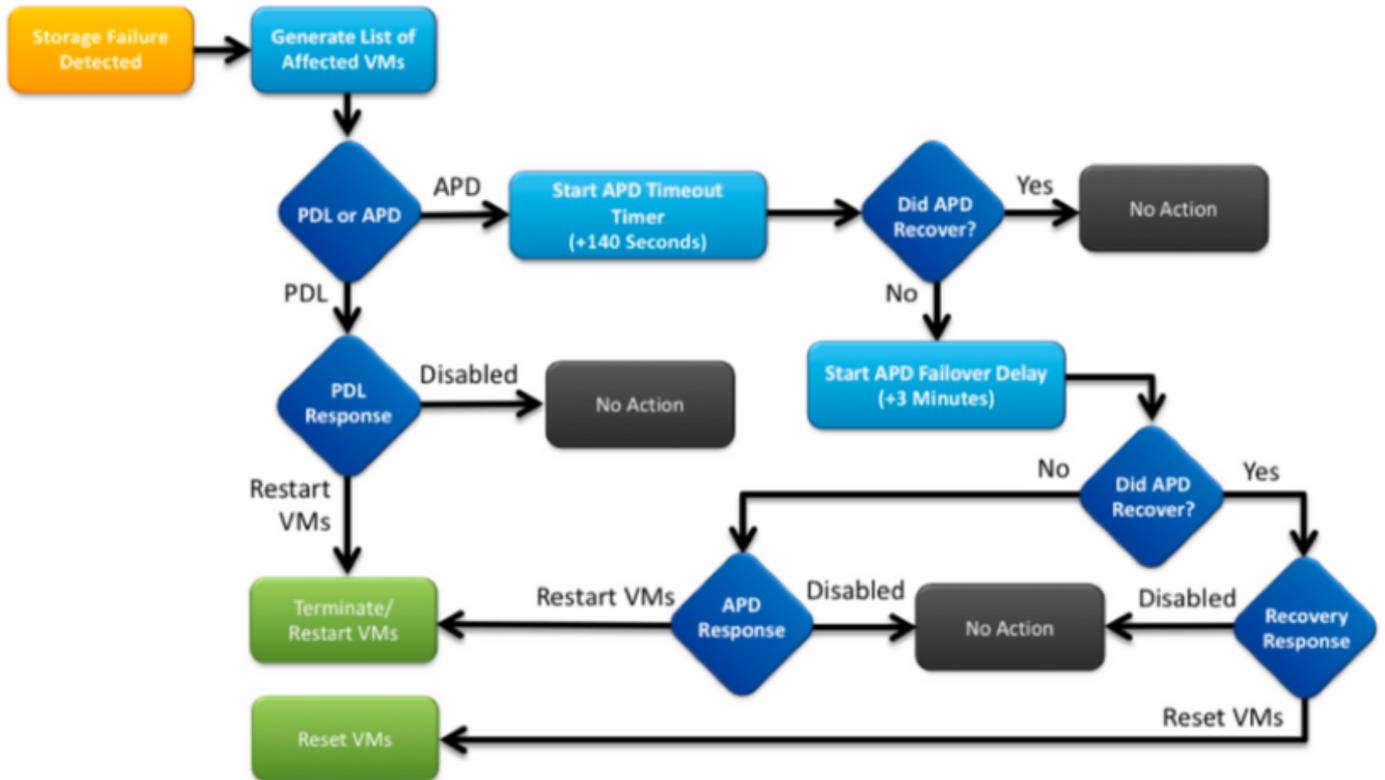
HXデータストアはESXIにどのようにマウントされますか。

- Hyperflexデータストアは、NFSマウントとしてESXiホストにマウントされます。NFSデータストアをマウントするには、NFSサーバIPが必要です。この場合は、eth1:0仮想フローティングインターフェイスです。
- Hyperflexクラスタは、仮想フローティングIP(eth0:mgmtip)とストレージデータ(eth1:0)の両方を使用し、各IPが1つの特定のストレージコントローラVM(StCtIVM)に割り振り当り当たります。これらは異なるStCtIVMになる場合があることに注意してください。
- 重要な点は、クラスタストレージデータIP(eth1:0)が、Hyperflexクラスタで作成されたデータストアのマウントに使用されるデータストアであることです。したがって、クラスタのすべてのノードから割り当てて到達可能にする必要があります。
- 現在eth1:0仮想IPを所有するStCtIVMに障害が発生した場合は、FHRP(First Hop Redundancy Protocol)と同様の方法で動作する別の使用可能なStCtIVMに「移行」する必要があります。

すべてのパスダウン

- APDは、ホストがストレージに到達できず、ストレージ・アレイから返されるPDL(Permanent Device Lost)SCSIコードがないことを意味します。
- 損失が一時的なものか否かが不明なため、ESXiホストはデフォルトで140以上(タイムアウト) + 3分(フェールオーバーの遅延)の通信を確立し続けます。ESXiホストは、ストレージデバイスに送信される非仮想マシンI/Oトラフィックに障害をします。
- vSphere 6.x環境でのAPDの詳細については、VMware KBの記事『[Permanent Device Loss \(PDL\) and All-Paths-Down \(APD\) in vSphere 5.x and 6.x](#)』を参照してください

• 次の図は、各断続的な段階について説明しています。



vCenterの一般的なエラーメッセージは次のとおりです。

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

問題の説明

ホストにAPDアラートが表示されたら、問題の説明を理解するために次の情報を取得します。

- 1つまたは複数またはすべてのホストに影響が及ぶ場合、および影響を受ける特定のホストがある場合
- 以前に変更があった場合（設定/アップグレードなど）
- 問題が最初に発生した時刻と問題が再発した場合のタイムスタンプ

トラブルシューティングワークフロー

APDをトラブルシューティングするには、vCenter、SCVM、およびESXiホストの3つのコンポーネントを調べる必要があります。

これらの手順は、観察されたAll Paths Down症状の原因を特定または絞り込むための推奨ワークフローです。この発注は慎重に行う必要はなく、お客様の環境で観察される特定の症状に従って適切に行う場合があります。

vCenter Serverのチェック：

vCenter Server(VCS)に接続し、影響を受けるホストに移動する

1. 関連オブジェクト -> 仮想マシンで、StCtIVMが起動して実行されていることを確認します
2. 関連オブジェクト -> Datastoresを選択し、NFSデータストアに「inaccessible」と表示されるかどうかを確認します。データストアがアクセス可能でステータスがあるような場合は、[Summary]タブでAPDイベントを「Reset to Green」にして、後でアラートがポップアップするかどうかを確認できます
3. [Monitor] -> [Issues] および[Monitor] -> [Events]は、APDが最初に検出されたタイミングに関する情報を提供します。

すべてのStCtIVMをチェックインします。

すべてのStCtIVMに接続し、次のポイントを確認します。この場合、[MobaXterm](#) 説明します。

1. dateまたはntpq -pを使用して、すべてのStCtIVMが同じ時刻になっていることを確認します。StCtIVM上の時間ずれは、Zookeeperデータベースの同期に関する問題を引き起こす可能性があるため、すべてのStCtIVM間で同期させることが重要です。ntpサーバのアスタリスク記号は、SCVMのNTPが同期されていることを示します。

```
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. アップグレード中にAPDが発生した場合は、どのStCtIVMが完全にアップグレードされていないかを確認し、特に最後に失敗したStCtIVMを特定します。これは、以前はeth1:0を保持していたデバイスであった可能性があります dpkg -lを使用して、| grep -i springpathを使用して、StCtIVMが完全にアップグレードされていないことを確認します。これは、バージョンが混在するspringpathパッケージが存在するためです。

```
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. 関連するすべてのサービスがservice_status.shを実行しているかどうかを確認します。主なサービスには、Springpath File System(storfs)、SCVM Client(scvmclient)、System Management Service(stMgr)、またはCluster IP Monitor(cip-monitor)があります。

```
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. これらのサービスまたは他の関連サービスのいずれかが起動していない場合は、**start** `<serviceName>`を使用して起動します(例：**start storfs**) service_status.shスクリプトを参照して、サービス名を取得できます(この例では、service_status.shスクリプトを使用します)。head -n25 /bin/service_status.shを実行し、サービスの実名を識別します。

```

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. ifconfig -aを使用して、ストレージクラスIP(eth1:0)を含むStCtlVMを識別します。StCtlVMにそのIPが含まれていない場合、storfsは1つ以上のノードで実行されていない可能性があります。

```

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. StCtlVMがCRMMasterに接続しているか、およびゾーキーパーサービスが起動して実行されているかを確認します **エコーSRVR | nc localhost 2181 and check mode is Leader, Follower or Standalone and if connections > 0**

```
root@help:~# echo srvr | 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
```

サービス出展者のステータスまたはps -ef | grep -i出展者

```
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
```

エラーまたはサービスが実行されていない場合は、次のログを確認し、zookeeperサービスを開始してください /var/log/springpath/exhibitor.log /var/log/springpath/stMgr.log サービス出展者がゾーキーパーサービスを開始する

7. すべてのStCtlVMからVCに到達可能かどうかを確認します STCLIクラスタ情報 | grep -i "url":VCのFQDNまたはIPを含むURLを表示します。ping <VC>を使用してVCへの接続を確認します。

```
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. FQDNを使用するクラスタの場合、DNSが到達可能かどうかを確認します stcli services dns showを使用して、StCtlVM上のDNS設定済みサーバを一覧表示します。ping <DNS_IP>およびhost <FQDN> <DNS_IP>を使用してDNSサーバへの接続と解決をテストします

```
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. すべてのStCtlVMに同じ量のiptablesエントリがあるかどうかを確認してください。 iptables -L | wc -l.不一致の場合は、TACケースをオープンしてください。

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

10. 現在のクラスタのステータスと状態は何ですか。 STCLIクラスタ情報 | lessまたはstcli cluster info | grep -i "active|state|unavailable"と表示される特定のノードを見つけようとすると、ストレージが使用できない状態になります。

```
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 --Outcall Billing Detail

```
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. どのデータストアがマウントされて利用可能ですか

```
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. stcliコマンドが長すぎるか、失敗した場合は、次のsysmtoolコマンドを試すことができます
 (stcliが動作している場合は使用しないでください) `sysmtool —ns cluster —cmd info sysmtool —ns cluster —cmd healthdetail sysmtool —ns datastore —cmd list`

StCtIVM:該当するESXiホストのStCtIVM

該当するESXiホストのStCtIVMに接続します

1. ストレージ・ クラスタIP(eth1:0)およびストレージ・ ネットワーク上の他のサーバ (StCtIVM上のeth1)への接続を確認します。
 stcliクラスタ情報の実行 | `grep -i -B 1 "stctl|hypervisor"`を使用して、クラスタに参加しているすべてのESXi Management IP、StCtIVM eth0 (管理) およびStCtIVM eth1 (ストレージデータ) を識別します接続`ping -l eth1 [-M do -s 8972] <target IP address>`をテストします。
ESXI VMK1とSCVM eth1の間でジャンボフレームをテストします。
2. それでも問題が特定されない場合は、次のログを調べてください `/var/log/springpath/debug-storfs.log` panics、seg faultまたはcritical events `grep -ai "segmentation|critical|panic" debug-storfs.log/var/log/springpath/stmgr.log``grep -i "oom|out of mem" /var/log/kern.log`
3. 最終的には、ノードのStCtIVMを再起動して問題が発生し、問題が解決しないかどうかを確認できます。

ESXiホストを確認します。

SSH経由で該当するESXiホストに接続し、次の操作を実行します。

1. `esxcli storage nfs list``Esxcfg-nas` -IれているNFS

```
[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
```

また、コマンド`cat /etc/vmware/esx.conf`を使用して、NFSマウントされたデータストアのESXi設定の一貫性を確認するために、`/etc/vmware/esx.conf`から確認することもできます | `grep -l nas`

2. `/var/log/vmkernel.log`を確認し、前の手順で特定したタイムスタンプに関する障害状態、マウントの問題、エラーの例を探します
3. IOVisor/NFS Proxy/SCVMClientのステータスを確認します コマンド`/etc/init.d/scvmclient status`を使用して、ESXiでサービスが実行されているかどうかを確認します [オプション `Jesxcli network ip connection list`を使用して、開いている接続があるかどうかを確認できます | `grep -i "proto|scvmclient"SCVMClient VIBがHXクラスタと同じバージョンであるかどうかを確認します。Esxcliソフトウェア VIBリスト | grep -i spring`

```
[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		

/var/log/scvmclient.logをチェックして、「clustermapを取得できません」というエラーが表示されるかどうかを確認します必要に応じて、etc/init.d/scvmclientrestartを使用してSCVMClientサービスを再起動できます

4. `vmk1ESXiIP eth1:0esxcfg-vmknic -l:vmk nicIPMTUvmkping -l vmk1 [-v -s 8972] -d <IP> TESXi`
5. **esxcli**ハードウェアプラットフォームは、StCtlVmの名前で使用されるサーバSNを取得し、特定のStCtlVMが実行されているホストを迅速に特定するのに役立ちます。