

# Fehlerbehebung bei Hyperflex Storage NFS-Problemen mit allen Pfaden (APD)

## Inhalt

[Einführung](#)

[Wie werden HX Data Stores auf der ESXI installiert?](#)

[Alle Pfade nach unten](#)

[Beschreibung des Problems](#)

[Fehlerbehebung](#)

[Prüfungen im vCenter-Server:](#)

[Checkt alle StCtlVMs ein:](#)

[StCtlVM: StCtlVM eines betroffenen ESXi-Hosts](#)

[Prüfungen im ESXi-Host:](#)

## Einführung

In diesem Dokument finden Sie schnelle Informationen und Problemlösungsschritte, die durchgeführt werden können, um die Problemursache zu ermitteln, wenn in vCenter, in das Hyperflex-Cluster integriert ist, die Fehlermeldung "NFS all pfs down" angezeigt wird.

## Wie werden HX Data Stores auf der ESXI installiert?

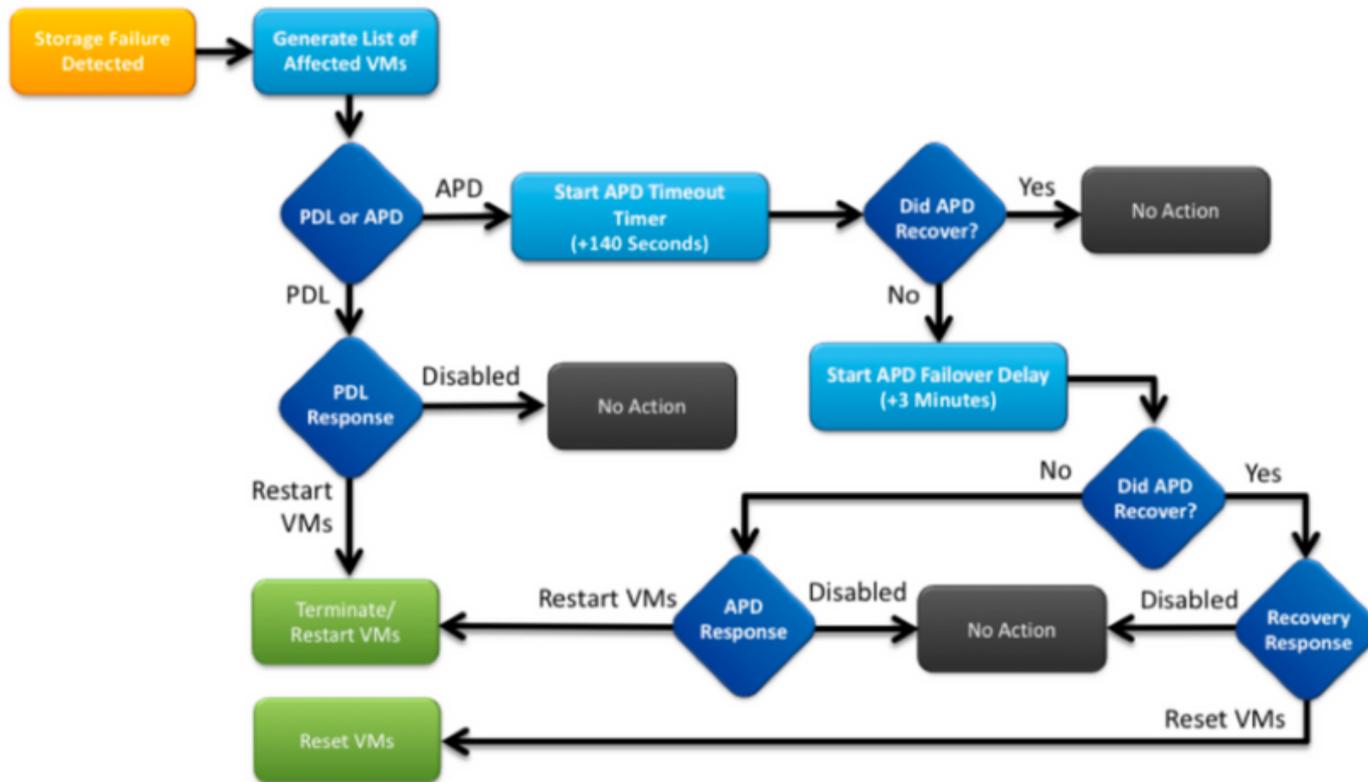
- HyperFlex-Datenspeicher werden als NFS-Anschlüsse auf den ESXi-Hosts bereitgestellt. Zur Bereitstellung eines NFS-Datenspeichers benötigen wir die NFS-Server-IP, die in unserem Fall die virtuelle Floating-Schnittstelle eth1:0 ist.
- HyperFlex-Cluster nutzt die Verwendung virtueller Floating-IP sowohl für Management (eth0:mgmtip) als auch für Storage-Daten (**eth1:0**), auf denen jede IP einem bestimmten Storage-Controller-VM (StCtlVM) zugewiesen wird. Beachten Sie, dass diese möglicherweise in verschiedenen StCtlVMs enden.
- Die Wichtigkeit ist, dass die Cluster-Storage-Daten-IP (eth1:0) verwendet wird, um die im **Hyperflex-Cluster** erstellten Datenspeicher(e) bereitzustellen. Daher ist es wichtig, dass sie **zugewiesen** und **von allen Knoten** des Clusters **erreichbar ist**.
- Bitte beachten Sie, dass bei einem Ausfall des StCtlVM, das derzeit die virtuelle eth1:0 IP besitzt, die Migration zu einem anderen verfügbaren StCtlVM erfolgen sollte, das ähnlich wie ein FHRP (First Hop Redundancy Protocol) funktioniert.

## Alle Pfade nach unten

- APD bedeutet, dass der Host den Speicher nicht erreichen kann und dass vom Speicher-Array kein Permanent Device Lost (PDL)-SCSI-Code zurückgegeben wird.
- Da er nicht weiß, ob der Ausfall temporär ist oder nicht, versucht er weiterhin, standardmäßig eine Kommunikation für mehr als 140 s (Timeout) + 3 Min. (Failover-Verzögerung) einzurichten. Der ESXi-Host beginnt, den E/A-Datenverkehr eines nicht virtuellen Systems

auszufallen, der an das Speichergerät gesendet wird.

- Weitere Informationen zu APD in vSphere 6.x-Umgebung finden Sie in dem VMware KB-Artikel [Permanent Device Loss \(PDL\) und All-Paths-Down \(APD\) in vSphere 5.x und 6.x](#).
- Im folgenden Bild werden die einzelnen Phasen dargestellt:



In vCenter wird eine typische Fehlermeldung angezeigt.

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

## Beschreibung des Problems

Wenn Sie APD-Warnungen auf Ihren Hosts sehen, erhalten Sie die folgenden Informationen, um die Problembeschreibung besser zu verstehen:

- Wenn ein/mehrere/alle Hosts betroffen waren und einige Hosts betroffen waren
- Wenn Änderungen zuvor durchgeführt wurden (Konfiguration/Upgrade/etc.)
- Der Zeitstempel, der angibt, wann das Problem erstmals beobachtet wurde und ob das Problem wiederholt auftritt.

## Fehlerbehebung

Zur Fehlerbehebung bei APD müssen drei Komponenten untersucht werden: vCenter, SCVM und ESXi-Host.

Diese Schritte sind ein empfohlener Workflow, um die Quelle des Symptoms All Paths Down (Alle

Pfade-Down) zu ermitteln oder einzugrenzen. Bitte beachten Sie, dass diese Bestellung nicht genau befolgt werden muss und Sie sie entsprechend den in der Kundenumgebung beobachteten besonderen Symptomen anpassen können.

## Prüfungen im vCenter-Server:

Herstellen einer Verbindung zum vCenter-Server (VCS) und Navigieren zu einem **betroffenen Host**

1. **Ähnliche Objekte -> Virtuelle Systeme** und Bestätigung der Inbetriebnahme von StCtIVM
2. **Ähnliche Objekte -> Datenspeicher** und bestätigen, ob NFS-Datenspeicher **"unzugänglich"** anzeigen. Wenn der **Zugriff** auf die Datenspeicher und der Status angezeigt werden, können Sie auf der Registerkarte **Zusammenfassung** das APD-Ereignis auf "Reset to Green" (Auf Grün zurücksetzen) zurücksetzen und später überprüfen, ob die Warnmeldung später erneut angezeigt wird.
3. **Überwachen -> Probleme** und **Überwachen -> Ereignisse** sollten Informationen darüber enthalten, wann der APD zum ersten Mal erkannt wurde.

## Checkt alle StCtIVMs ein:

Stellen Sie eine Verbindung zu **allen StCtIVMs** her, und überprüfen Sie die folgenden Zeiger. Sie können [MobaXterm](#) Software.

1. Überprüfen Sie, ob alle StCtIVMs die gleiche Zeit mit **date** oder **ntpq -p** haben. Zeitverschiebungen auf StCtIVM können zu Problemen bei der Synchronisierung der Zookeeper-Datenbank führen. Daher ist es von entscheidender Bedeutung, dass diese Daten zwischen allen StCtIVMs synchronisiert werden. Der Signaleingang des NTP-Servers gibt an, dass das NTP des SCVM synchronisiert wird.

```
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. Wenn APD während eines **Upgrades** auftrat, sollten Sie prüfen, welche **StCtIVMs nicht vollständig aktualisiert** wurden, **und insbesondere ermitteln, welches System zuletzt fehlschlug**. Es ist möglich, dass es derjenige war, der zuvor das eth1:0 hielt. Verwenden Sie **dpkg -l | grep -i springpath** zur Identifizierung der StCtIVMs, die nicht vollständig aktualisiert wurden, da sie gemischte Version springpath-Pakete haben werden.

```
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. Überprüfen Sie, ob alle relevanten Services ***service\_status.sh*** ausführen: Zu den wichtigsten Services gehören das Springpath File System (*storfs*), der SCVM Client (*scvmclient*), der System Management Service (*stMgr*) oder der 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. Wenn einer dieser oder andere relevante Services nicht verfügbar ist, starten Sie ihn ***mithilfe vonStart <ServiceName>***. z. B.: ***start storfs*** Sie können das Skript `service_status.sh` verwenden, um die Dienstnamen zu erhalten. Führen Sie einen **Kopf -n25 /bin/service\_status.sh** und identifizieren Sie den echten Service-Namen.

```

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. Identifizieren Sie, **welches StCtlVM das StorageCluster IP (eth1:0) enthält**, indem Sie **ifconfig -a verwenden**. Wenn keine StCtlVM diese IP enthält, werden die Storfs möglicherweise nicht auf einem oder mehreren Knoten ausgeführt.

```
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. Überprüfen Sie, ob **StCtlVM mit CRMMaster in Kontakt** ist und ob der **Zookeeper-Service in Betrieb ist**. **Echo-Server | nc localhost 2181** und prüfen, ob der Modus Leader, Follower oder Standalone ist und ob Verbindungen > 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
```

#### **Ausstellerstatus oder ps-ef | grep -i Aussteller**

```
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
```

Falls ein Fehler oder Dienst nicht ausgeführt wird, können Sie die folgenden Protokolle überprüfen und versuchen, den Zookeeper-Dienst zu starten **/var/log/springpath/exhibitor.log** und **/var/log/springpath/stMgr.log****Service Aussteller starten Zookeeper Service**

7. Überprüfen Sie, ob **VC von allen StCtlVMs erreichbar** ist. **stcli cluster info | grep -i "url"**, um die URL anzuzeigen, die entweder FQDN oder IP von VC enthält. Überprüfen Sie die Verbindung zu VC mit **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. Bestätigen Sie, ob **DNS im Fall eines Clusters mit FQDN erreichbar** ist. **stcli services dns zeigen** die DNS-konfigurierten Server auf StCtlVM an. Testen der **Konnektivität** und **Auflösung** zu **DNS-Servern** mit **ping <DNS\_IP>** und **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. Bestätigen Sie, dass alle **StCtlVMs die gleiche Anzahl von iptables-Einträgen** haben: **iptables -L | wc -l**. Falls sie nicht übereinstimmen, öffnen Sie bitte ein TAC-Ticket.

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

10. Wie lauten der aktuelle Cluster-Status und Status? **stcli cluster info | weniger** oder **stcli Clusterinformationen | grep -i "active|state|unavailable"**, wenn versucht wird, herauszufinden, welche Knoten bei nicht verfügbarem Speicher angezeigt werden.

```
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 —Details zum Abrufen der Speicherclusterdetails**

```
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. Welche Datenspeicher sind bereitgestellt und verfügbar?**

```
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. Falls stcli-Befehle zu lange dauern oder fehlschlagen, können Sie die folgenden sysmtool-Befehle verwenden (nicht verwenden, wenn stcli funktioniert) **sysmtool —ns-Cluster —cmd info sysmtool —ns Cluster —cmd gesundheitsdetail sysmtool —ns datastore —cmd-Liste**

## StCtIVM: StCtIVM eines betroffenen ESXi-Hosts

Herstellen einer Verbindung zum StCtIVM des betroffenen ESXi-Hosts

1. Überprüfen Sie die Verbindung zur Storage-Cluster-IP (eth1:0) und zu anderen Servern im Storage-Netzwerk (**eth1** auf StCtIVMs).  
Ausführen **von stcli cluster info | grep -i -B 1 "stctl\|hypervisor"** zur Identifizierung aller am Cluster teilnehmenden ESXi Management-IP-, StCtIVM eth0 (Mgmt)- und StCtIVM eth1- (Storage-Daten) Testen Sie die Verbindung **ping -I eth1 [-M do -s 8972] <Ziel-IP-Adresse>**, **Jumbo-Frames-Test zwischen ESXI VMK1 und SCVM eth1**.
2. Wenn das Problem immer noch nicht identifiziert wurde, können Sie sich die folgenden Protokolle ansehen. **/var/log/springpath/debug-storfs.log** Prüfen Sie, ob Panik, Fehler oder kritische Ereignisse **grep -ai "Segmentation\|critical\|panic" debug-storfs.log/var/log/springpath/stmgr.log** Überprüfen Sie, ob das Speicherproblem **grep -i "oom\|out of mem" ist /var/log/kern.log**
3. Letztlich können Sie versuchen, das **StCtIVM des Knotens** neu zu starten, bei dem das **Problem besteht, und überprüfen, ob das Problem weiterhin besteht.**

## Prüfungen im ESXi-Host:

Stellen Sie über **SSH** eine Verbindung zu einem betroffenen **ESXi-Host** her, und führen Sie die folgenden Schritte durch:

1. **esxcli storage nfs list** oder **Esxcfg-nas -l**, UM die aktuell installierten NFS-Datenspeicher aufzulisten und ob sie zugänglich sind

```

[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

```

Sie können die Konsistenz der ESXi-Konfiguration in den NFS-gemounteten Datenspeichern

auch von `/etc/vmware/esx.conf` bestätigen. Verwenden Sie hierzu den Befehl `cat /etc/vmware/esx.conf`. | `grep -l nas`

- Überprüfen Sie `/var/log/vmkernel.log`, und suchen Sie nach Fehlfunktionen, Bereitstellungsproblemen oder Fehlern im Zeitstempel, der in den vorherigen Schritten angegeben wurde.
- Überprüfen Sie den Status von IOVisor/NFS Proxy/SCVMClient. Überprüfen Sie, ob der **Dienst** unter dem Befehl `/etc/init.d/scvmclient-Status` auf ESXi ausgeführt wird. [Optional] Sie können mithilfe der `esxcli netzwerk ip connection list` überprüfen, ob offene Verbindungen vorhanden sind. | `grep -i "proto|scvmclient"` Bestätigen Sie, ob SCVMClient VIB die gleiche Version wie Ihr HX-Cluster ist, `Esxcli-Software-vib-Liste`. | `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
```

Überprüfen Sie `/var/log/scvmclient.log`, ob Fehler vorliegen, nämlich "nicht in der Lage, ClusterMap zu erhalten". Sie können den SCVMClient-Dienst bei Bedarf über `etc/init.d/scvmclientrestart` neu starten.

- Überprüfen der Netzwerkverbindung mit anderen ESXi-Hosts im vmk1-Netzwerk, insbesondere mit dem Speicher-Cluster IP `eth1:0esxcfg-vmknic -l` zum Abrufen von Informationen über die vmk nic-Details, z. B. IP, Maske und MTU `vmkping -l vmk1 [-v -s 8972] -d <Ziel-IP-Adresse>` test-Anbindung [optional mit Jumbo Frames] zwischen ESXi-Hosts im Controller-Datennetzwerk
- esxcli Hardwareplattform erhalten** die Server-SN, die auf dem Namen des StCtIVm verwendet wird und Ihnen helfen kann, schnell herauszufinden, auf welchem Host ein bestimmtes StCtIVM ausgeführt wird.