Disk drive troubleshooting

CVM internal storage, SATADOM, and M.2 drives

Disk drive troubleshooting overview

Disk drive issues include host boot disk and CVM storage disk problems. The following tools can be used to diagnose disk issues:

- Prism
- SMART utilities
- NCC CLI commands
- SATADOM diagnostics via hypervisor commands

This section also include sample scenarios that help to identify of disk issues.



Checking disk status in Prism

In Prism, users can check for disk status in the following two places:

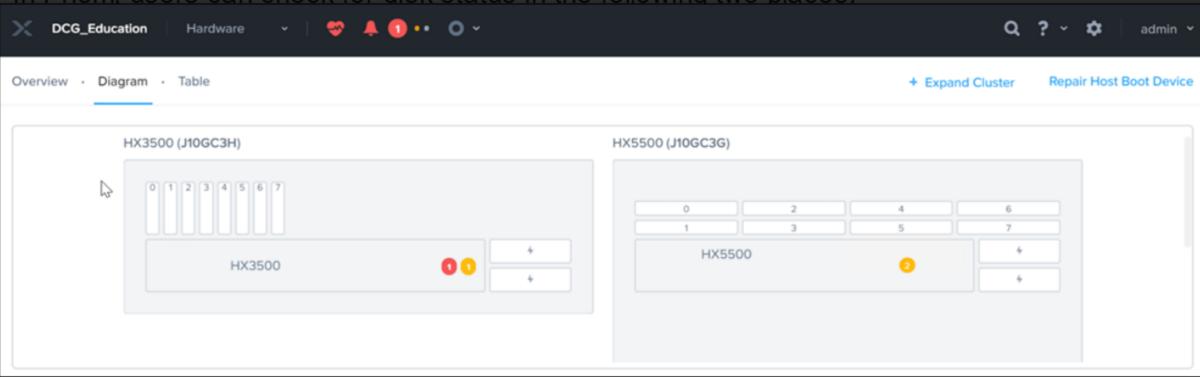
- Go to Hardware → Diagram to check the status of the disks in each node.
 (click HERE to see the image)
- Go to Hardware → Table to check the summary and status of each disk in each node.
 (click HERE to see the image)



×

Checking disk status in Prism

In Prism. users can check for disk status in the following two places:







packing dick statue in Driem ♥ A 0 · · O · Q ? - # DCG_Education Hardware admin ~ Overview · Diagram · Table + Expand Cluster Repair Host Boot Device • Disk Host Switch 1-10 of 36 · < > · · · search in table Q · Disk ID Hypervisor IP Status Disk Usage Disk IOPS Disk IO B/W Disk Avg IO Latency Serial Number Tier 26 HDD Online 181.3 MiB of 801.23 GiB 0 0 KBps 9XG9XEAX 10.10.3.14 0 ms 28 HDD 0 0 KBps 9XG9XELC 10.10.3.14 Online 99.19 MiB of 841.23 GiB 0 ms 29 9XG9YAZG 10.10.3.14 HDD 0 0 KBps Online 163.69 MiB of 841.23 GiB 0 ms 30 10.10.3.14 HDD 181.54 MIB of 841.23 GIB 7 KBps 9XG9XF7X Online 0 14.54 ms 31 9XG9XFH3 10.10.3.14 HDD Online 162.04 MiB of 841.23 GiB 0 0 KBps 3.55 ms BTHV608203CY40 32 10.10.3.14 SSD 9.31 GIB of 215.74 GIB 12 175 KBps 0.13 ms Online ONGN BTHV6082032H40 33 9.42 GIB of 215.74 GIB 10.10.3.14 SSD 18 198 KBps Online 0.3 ms ONGN Z4H06FRM0000R 35 10.10.3.13 HDD 0 4 KBps 12.19 ms Online 1.16 GIB of 1.65 TIB 623NVHW



SMART utilities

For further, in-depth disk analysis, users can use SMART utilities.

Self-Monitoring, Analysis, and Reporting Technology (SMART) is a monitoring system included in computer HDDs, SSDs, and eMMC drives. Its primary function is to detect and report indicators of drive reliability in order to anticipate imminent hardware failures. When SMART data indicates a possible imminent drive failure, software running on the host system can notify the user so preventative action can be taken to prevent data loss. The failing drive can also be replaced and data integrity maintained.

To access SMART utilities, log in to CVM and manually run smartctl.

- 1. #ls -la /dev/sd* to list all disks in the system
- 2. #sudo smartctl -x /dev/sd<x> -T permissive smartctl can also be invoked to run a short or a long test in the background

```
#sudo smartctl /dev/sd<x> -t short
#sudo smartctl /dev/sd<x> -t long
```

The status of the test can be checked with the #sudo smartctl /dev/sd<x> -a command

Go to the next page for an example.



Lenovo

Example of smartctl test output

```
=== START OF INFORMATION SECTION ===
                                     00YC341 00YC344LEN
Device Model:
                 SSDSC2BA400G4N
Serial Number: BTHV608204B2400NGN
LU WWN Device Id: 5 5cd2e4 04c1c9c82
Firmware Version: G201LS36
User Capacity: 400,088,457,216 bytes [400 GB]
Sector Sizes: 512 bytes logical, 4096 bytes physical
                                                                                      Disk information
Rotation Rate: Solid State Device
Form Factor: 2.5 inches
Device is: Not in smartctl database [for details use: -P showall]
ATA Version is: ACS-2 T13/2015-D revision 3
SATA Version is: SATA 2.6, 6.0 Gb/s (current: 6.0 Gb/s)
Local Time is:
               Mon Sep 16 22:32:52 2019 EDT
SMART support is: Available - device has SMART capability.
SMART support is: Enabled
=== START OF READ SMART DATA SECTION ===
SMART overall-health self-assessment test result: PASSED
SMART Error Log Version: 1
No Errors Logged
                                                                                                        SMART test results
SMART Self-test log structure revision number 1
    Test Description
                       Status
                                               Remaining LifeTime (hours) LBA of first error
    Short offline
                  Completed without error
                                                            27562
    Short offline
                  Completed without error
                                                            20111
                      Completed without error
# 3 Extended offline
                                                     00%
                                                            20109
```

Host boot disk diagnostics – SATADOM

To troubleshoot SATADOM issue individually, the following methods can be utilized. Before proceeding, users should first ensure the server firmware is up to date by checking the Lenovo HX best recipe Web site at https://support.lenovo.com/tw/en/solutions/HT508092.

- Run CLI command:
 - Disk firmware information can be found using the #ncc hardware_info show_hardware_info or smartctl /dev/sd<x> -a commands.
 - Check the health of the SATADOM with the #ncc health_checks hardware_checks disk checks run all command.
- The host boot disk health status can also checked in different hypervisors (ESXi, Hyper-V, and AHV).
- If the SATADOM is not detected, turn off the node and wait 30 to 40 seconds. After turning the node back on, users should be able to see the SATADOM. If so, upgrade the firmware. If the SATADOM is still not detected, replace it with a new one.
- If users see I/O errors on the CVM console, the host hangs (most often on AHV), or errors on ESXi, turn off the node. Wait 30 to 40 seconds, and then check the SATADOM status. If the SATADOM status is still abnormal, replace the node and perform a firmware upgrade.



ESXi SATADOM CLI check – example

SATADOM2DSL 3ME

Scroll down for more information.

- Use PuTTY to ssh log in to the ESXi console IP. This is the name of the disk. Check the SATADOM disk list. #ls /dev/disks | grep -v ":" | grep -i satadom 20151021AA8511027060 t10.ATA
- 3. Check the status of the SATADOM. The health status can be checked with following commands:

```
#localcli storage core device smart get -d <satadom name>
#./iSMART -d <satadom name>
```

Example:

Lenovo

```
#localcli storage core device smart get -d t10.
ATA InnoLite SATADOM D150QV2DL 20131009AA5500000077
Parameter Value Threshold Worst
Health Status OK N/A N/A
Media Wearout Indicator N/A N/A
Write Error Count N/A N/A N/A
Read Error Count 100 50 100
Power-on Hours 100 0 100
Power Cycle Count 100 0 100
Reallocated Sector Count 100 50 100
```

ESXi SATADOM CLI check – example

Scroll down for more information.

Media Wearout Indicator N/A N/A N/A
Write Error Count N/A N/A N/A
Read Error Count 100 50 100
Power-on Hours 100 0 100
Power Cycle Count 100 0 100
Reallocated Sector Count 100 50 100
Raw Read Error Rate 100 50 100
Drive Temperature 45 0 100
Driver Rated Max Temperature N/A N/A
Write Sectors TOT Count N/A N/A N/A
Read Sectors TOT Count N/A N/A N/A
Initial Bad Block Count N/A N/A N/A

To list only the health status of the SATADOM without first checking its name, you can run the following command:

```
#1s /dev/disks | grep -v ":" | grep -i satadom | while read name; do localcli storage core device smart get -d $name | grep -B2 "Health Status"; done
Parameter Value Threshold Worst
```

Health Status OK N/A N/A



Hyper-V SATADOM CLI check

Log in to Hyper-V via a remote desktop or with the winsh command, and then issue these two commands:

```
    #Get-WMIObject -Class Win32_DiskDrive | Format-List *
    #Get-WMIObject -Class Win32 DiskDrive | Select Caption, Status
```

Example:

```
#winsh "Get-WMIObject -Class Win32_DiskDrive | Select Caption, Status"
Caption Status
------
ATA INTEL SSDSC2BX48 SCSI Disk Device OK
ATA ST2000NM0033-9ZM SCSI Disk Device OK
ATA ST2000NM0033-9ZM SCSI Disk Device OK
SATADOM-SL 3ME OK
```



CVM SATADOM CLI check

- Use PuTTY to ssh into Nutanix CVM.
- Use the following command to check the SATADOM device name:

```
# ls /dev/sd*
/dev/sda /dev/sda1
```

Use the following smartctl command to check the status:

```
# sudo smartctl -a /dev/sda | grep result
SMART overall-health self-assessment test result: PASSED
```

This is the same as the command used to check physical SSD and HDD status from the CVM.



Host boot disk diagnostics – M.2 drive

Troubleshooting issues with M.2 boot drives is similar to troubleshooting SATADOM issues.

- For CVM, troubleshooting M.2 drives is identical to troubleshooting an SSD on a CVM.
- ESXi utilizes a different set of commands.
 - To view connected devices:

```
#esxcfq-scsidevs -l | grep '^t10'
```

To view diagnostic information:

```
#esxcli storage core device smart get -d <disk_name>
```

• To check which M.2 drive is used as the boot drive, run the following command:

```
#esxcfg-info -s | grep -A10 "Diagnostic Partition"
```



Checking RAID volumes in CVM

Use the following command to check boot RAID in CVM:



CVM storage disk identification – list_disks

The list_disks command lists the disks attached to the system.

Note: Slot 0 is for media devices, such as a CDROM. It can be ignored.



CVM storage disk identification – Isscsi

The lsscsi command queries the LSI controller directly for information on connected disks.

```
#lsscsi
[0:0:0:0] cd/dvd NECVMWar VMware IDE CDR00 1.00 /dev/sr0
[2:0:0:0] disk ATA INTEL SSDSC2BA40 0265 /dev/sda
[2:0:1:0] disk ATA INTEL SSDSC2BA40 0265 /dev/sdb
[2:0:2:0] disk ATA ST91000640NS SN03 /dev/sdc
[2:0:3:0] disk ATA ST91000640NS SN03 /dev/sdd
[2:0:4:0] disk ATA ST91000640NS SN03 /dev/sde
[2:0:5:0] disk ATA ST91000640NS SN03 /dev/sdf
```



CVM storage disk identification – Isblk

The lsblk command shows how each block device is mapped to the Linux mount points on the CVM. The partitions making up the CVM RAID partitions can be seen here.

admin@NT	NX-J10G	C3G-	-A-CVM:	10.1	10.3.1	6:~\$ lsblk
NAME	MAJ:MIN	RM	SIZE	RO	TYPE	MOUNTPOINT
sda	8:0	0	372.6G	0	disk	
-sdal	8:1	0	10G	0	part	
L_md0	9:0	0	10G	0	raidl	
-sda2	8:2	0	10G	0	part	
L_md1	9:1	0	10G	0	raidl	
-sda3	8:3	0	40G	0	part	
L-md2	9:2	0	40G	0	raidl	/home
-sda4	8:4	0	312.6G	0	part	/home/nutanix/data/stargate-storage/disks/BTHV608204B2400NGN
sdb	8:16	0	372.6G	0	disk	
-sdb1	8:17	0	10G	0	part	
-md0	9:0	0	10G	0	raidl	
-sdb2	8:18	0	10G		part	
-md1	9:1	0	10G		raidl	
-sdb3	8:19	0	40G		part	
∟md2	9:2	0	40G	0	raidl	/home
-sdb4	8:20		312.6G		part	/home/nutanix/data/stargate-storage/disks/BTHV60820322400NGN
adc	8:32	0			disk	
-sdc1	8:33	0	1.8T		part	/home/nutanix/data/stargate-storage/disks/Z4H07ENW0000R623R3XP
sdd	8:48	0	1.8T		disk	
-sdd1	8:49	0	1.8T		part	/home/nutanix/data/stargate-storage/disks/Z4H06FRM0000R623NVHW
sde	8:64	0			disk	
-sde1	8:65	0			part	/home/nutanix/data/stargate-storage/disks/Z4H06FQ20000R616N0J9
sdf	8:80	0			disk	
-sdf1	8:81	0			part	/home/nutanix/data/stargate-storage/disks/Z4H07R4Q0000R627HDM0
adg	8:96	0	1.8T		disk	
-sdg1	8:97	0	1.8T		part	/home/nutanix/data/stargate-storage/disks/Z4H06FMF0000R623R0CX
sdh	8:112	0			disk	
-sdh1	8:113	0	1.8T		part	/home/nutanix/data/stargate-storage/disks/Z4H07EMV0000R623R36K
sr0	11:0	1	8.3M		rom	According to
loop0	7:0	0	256M		loop	/var/tmp
admin@NT	NX-J10G	USG-	-A-CVM:	LU.	10.3.1	6:~\$



CVM storage disk identification – Isiutil

lsiutil: LSI provides a utility which returns SCSI information for the attached devices and allows access to protocol level count.

```
#sudo /home/nutanix/cluster/lib/lsi-sas/lsiutil -a 12,0,0 16
(Click HERE to see the output.)
#sudo /home/nutanix/cluster/lib/lsi-sas/lsiutil -a 12,0,0 20
(Click HERE to see the optput.)
```

Note: To get disks and their location relationship, use the <code>list_disks</code> and <code>lsiutil</code> commands.

list_disks	/home/nutanix/cluster/lib/lsi-sas/lsiutil -a 12,0,0 20					
Slot Disk Model Serial Size	Adapter Phy 0: Link Up, No Errors → /dev/sda					
0	Adapter Phy 1: Link Up, No Errors > /dev/sdb					
1 /dev/sda INTEL SSDSC2BA40 BTTV324500SX400HGN 400 GB	Adapter Phy 2: Link Up, No Errors > /dev/sdc					
2 /dev/sdb INTEL SSDSC2BA40 BTTV324301ZB400HGN 400 GB	Adapter Phy 3: Link Up, No Errors > /dev/sdd					
3 /dev/sdc ST91000640NS 9XG5GHQR 1.0 TB	Adapter Phy 4: Link Up, No Errors > /dev/sde					
4 /dev/sdd ST91000640NS 9XG4NW6Q 1.0 TB	Adapter Phy 5: Link Up, No Errors > /dev/sdf					
5 /dev/sde ST91000640NS 9XG4PH3P 1.0 TB	Adapter Phy 6: Link Up, No Errors -> /dev/sdg					
6 /dev/sdf ST91000640NS 9XG4PE7H 1.0 TB	Adapter Phy 7: Link Down, No Errors → no device					



Disk drive troubleshoo

MPT Port found





CVM stor

lsiutil: LS

allows acces

(Click HERE

(Click HERE

Note: To get

Port Name Chip Vendor/Type/Rev MPT Rev Firmware Rev IOC 1. ioc0 LSI Logic SAS3008 CO 205 0f000200

LSI Logic MPT Configuration Utility, Version 1.70, July 30, 2013

SAS3008's links are 12.0 G, 12.0 G, 6.0 G, 6.0 G, 12.0 G, 12.0 G, 12.0 G, 12.0 G

В	T	SASAddress	PhyNum	Handle	Parent	Type		
		500605b00ac41c10		0001		SAS Initiator		
		500605b00ac41c11		0002		SAS Initiator		
		500605b00ac41c12		0003		SAS Initiator		
		500605b00ac41c13		0004		SAS Initiator		
		500605b00ac41c14		0005		SAS Initiator		
		500605b00ac41c15		0006		SAS Initiator		
		500605b00ac41c16		0007		SAS Initiator		
		500605b00ac41c17		8000		SAS Initiator		
0	3	5000c50084ddfe79	0	0009	0001	SAS Target		
0	1	4433221103000000	3	000a	0002	SATA Target		
0	0	4433221102000000	2	000ь	0003	SATA Target		
0	2	5000c50084ddfa99	1	000c	0004	SAS Target		
0	7	5000c50084ddf9cd	4	D000d	0005	SAS Target		
0	6	5000c50084de07c5	5	000e	0006	SAS Target		
0	4	5000c50084de01f9	6	000f	0007	SAS Target		
0	5	5000c50084f2b1c5	7	0010	8000	SAS Target		

	** ***	mt	** ***					
Type	NumPhys	PhyNum	Handle		PhyNum	Handle	Port	Speed
Adapter	8	0	0001	>	0	0009	0	12.0
		1	0004	>	0	000c	3	12.0
		2	0003	>	0	000b	2	6.0
315		3	0002	>	0	000a	1	6.0
511		4	0005	>	0	D000	4	12.0
57		5	0006	>	0	000e	5	12.0
		6	0007	>	0	000f	6	12.0
		7	0008	>	0	0010	7	12.0

vices and

1 commands.

```
no device
```

Enclosure Handle Slots SASAddress B T (SEP) 500605b00ac41c10 0001

×

Lenovo

Lenovo

Offline disk example

A disk might be completely dead and offline. In this example, the disk in Slot 2 is not detected (the disk in Slot 0 can be ignored) even though /dev/sdb is expected to be online. With the listutil command, you can see that the link is down (Adapter Phy 1: Link Down, No Errors). Replace the disk.

#list disks Slot Disk Model Serial Size 1 /dev/sda INTEL SSDSC2BA40 BTTV324500SX400HGN 400 GB /dev/sdb is missing /dev/sdc ST91000640NS 9XG4NW60 1.0 TB 4 /dev/sdd ST91000640NS 9XG4NW60 1.0 TB 5 /dev/sde ST91000640NS 9XG4PH3P 1.0 TB 6 /dev/sdf ST91000640NS 9XG4PE7H 1.0 TB 7 /dev/sdq ST91000640NS 9XG4PA9H 1.0 TB #sudo /home/nutanix/cluster/lib/lsi-sas/lsiutil -a 12,0,0 20 LSI Logic MPT Configuration Utility, Version 1.70, July 30, 2013 1 MPT Port found Port Name Chip Vendor/Type/Rev MPT Rev Firmware Rev IOC

Offline disk example

A disk might be completely dead and offline. In this example, the disk in Slot 2 is not detected (the disk in Slot 0 can be ignored) even though /dev/sdb is expected to be online. With the listutil command, you can see that the link is down (Adapter Phy 1: Link Down, No Errors). Replace the disk.

```
#SUGO /nome/nutanix/cluster/lib/isi-sas/isiutil -a iz,U,U ZU
LSI Logic MPT Configuration Utility, Version 1.70, July 30, 2013
1 MPT Port found
Port Name Chip Vendor/Type/Rev MPT Rev Firmware Rev IOC
1. ioc0 LSI Logic SAS2008 B2 200 0e000000 0
Diagnostics menu, select an option: [1-99 or e/p/w or 0 to quit] 12
Adapter Phy 0: Link Up, No Errors
                                                                  /dev/sdb link is down
Adapter Phy 1: Link Down, No Errors
Adapter Phy 2: Link Up, No Errors
Adapter Phy 3: Link Up, No Errors
Adapter Phy 4: Link Up, No Errors
Adapter Phy 5: Link Up, No Errors
Adapter Phy 6: Link Up, No Errors
Adapter Phy 7: Link Down, No Errors
```

Disk online, but marked as bad

A disk could be reported as online (for example, the disk in Slot 2) but marked as bad by stargate/hades. In <code>lsiutil</code>, it can be seen that the error counter is increasing (<code>Adapter Phy 1</code>). This might be correlated with issues with the chassis or the LSI controller on the node. If multiple ports report incrementing error counts, the issue might be with the LSI controller itself or with the wiring between the node and the disks. It might be necessary to replace either the node or the chassis.

```
#sudo /home/nutanix/cluster/lib/lsi-sas/lsiutil -a 12,0,0 20
LSI Logic MPT Configuration Utility, Version 1.70, July 30, 2013
1 MPT Port found
Port Name Chip Vendor/Type/Rev MPT Rev Firmware Rev IOC
1. ioc0 LSI Logic SAS2008 B2 200 0e000000 0
Diagnostics menu, select an option: [1-99 or e/p/w or 0 to quit] 12
Adapter Phy 0: Link Up, No Errors
Adapter Phy 1: Link Up, 9827532
Adapter Phy 2: Link Up, No Errors
Adapter Phy 3: Link Up, No Errors
Adapter Phy 4: Link Up, No Errors
Adapter Phy 5: Link Up, No Errors
Adapter Phy 6: Link Up, No Errors
Adapter Phy 7: Link Down, No Errors
```

