



Monitor and Manage Express Flash NVMe PCIe SSD in VMware ESXi

Karan Singh Gandhi
Arunava Das Gupta
Krishnaprasad K

July 2015

Revisions (required)

Date	Description
July 2015	Initial release

THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

© 2015 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell.

PRODUCT WARRANTIES APPLICABLE TO THE DELL PRODUCTS DESCRIBED IN THIS DOCUMENT MAY BE FOUND AT: <http://www.dell.com/learn/us/en/19/terms-of-sale-commercial-and-public-sector>

Trademarks used in this text:

Dell™, the Dell logo, Dell Boomi™, Dell Precision™, OptiPlex™, Latitude™, PowerEdge™, PowerVault™, PowerConnect™, OpenManage™, EqualLogic™, Compellent™, KACE™, FlexAddress™, Force10™ and Vostro™ are trademarks of Dell Inc. Other Dell trademarks may be used in this document. Cisco Nexus®, Cisco MDS®, Cisco NX-OS®, and other Cisco Catalyst® are registered trademarks of Cisco System Inc. EMC VNX®, and EMC Unisphere® are registered trademarks of EMC Corporation. Intel®, Pentium®, Xeon®, Core® and Celeron® are registered trademarks of Intel Corporation in the U.S. and other countries. AMD® is a registered trademark and AMD Opteron™, AMD Phenom™ and AMD Sempron™ are trademarks of Advanced Micro Devices, Inc. Microsoft®, Windows®, Windows Server®, Internet Explorer®, MS-DOS®, Windows Vista® and Active Directory® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. Red Hat® and Red Hat® Enterprise Linux® are registered trademarks of Red Hat, Inc. in the United States and/or other countries. Novell® and SUSE® are registered trademarks of Novell Inc. in the United States and other countries. Oracle® is a registered trademark of Oracle Corporation and/or its affiliates. Citrix®, Xen®, XenServer® and XenMotion® are either registered trademarks or trademarks of Citrix Systems, Inc. in the United States and/or other countries. VMware®, Virtual SMP®, vMotion®, vCenter® and vSphere® are registered trademarks or trademarks of VMware, Inc. in the United States or other countries. IBM® is a registered trademark of International Business Machines Corporation.

Broadcom® and NetXtreme® are registered trademarks of Broadcom Corporation. Qlogic is a registered trademark of QLogic Corporation. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and/or names or their products and are the property of their respective owners. Dell disclaims proprietary interest in the marks and names of others.

Table of Contents

Introduction	5
Audience and Scope.....	5
What's NVMe?	5
Manage Express Flash NVMe PCIe SSD via CLI	5
Manage Express Flash NVMe PCIe SSD via Webclient.....	8
Manage Express Flash NVMe PCIe SSD using Dell Openmanage Storage Services	11
Significant Enumerated Properties of NVMe Devices:	11
Supported Operations:	12
Monitoring:	13
Conclusion.....	14
References	14

Figure 1 : lspci grep NVMe	6
Figure 2 : esxcfg-scsidevs -a	6
Figure 3 : esxcfg-scsidevs -c	6
Figure 4 : esxcfg-scsidevs -l -d	6
Figure 5 : esxcfg-module -g nvme	6
Figure 6 : esxcfg-module -i nvme	7
Figure 7 : esxcli storage core adapter stats	8
Figure 8 : esxcli storage core device list.....	8
Figure 9 : Properties of NVMe device(s)	9
Figure 10 : Formatting NVMe device as a VMFS datastore	10
Figure 11 : NVMe device as a virtual flash.....	11
Figure 12 : Dell OpenManage UI displaying NVMe device(s)	12
Figure 13 : OpenManage available tasks for NVMe device(s) of 2.5" form factor	13
Figure 14 : OpenManage available tasks for NVMe device(s) of add-in card form factor	13

Introduction

This document provides an insight into monitoring and managing Dell PowerEdge Express Flash NVMe PCIe SSD devices connected to VMware ESXi. There are in-built options available to monitor and manage certain aspects of Express Flash NVMe PCIe SSD in ESXi.

From VMware ESXi 6.0 onwards, Dell OpenManage version 8.1.0 also supports monitoring and managing Express Flash NVMe PCIe SSD devices. Dell OpenManage for VMware ESXi supporting Express Flash NVMe PCIe SSD can be downloaded from [Dell support](#). This article briefly touches upon the options available for the end users to monitor and manage Express Flash NVMe PCIe SSD in VMware ESXi.

Audience and Scope

The scope of this article is to list out all the possible options for monitoring and managing Express Flash NVMe PCIe SSD devices in VMware ESXi environment. This paper might be useful for NVMe end users and datacenter administrators to monitor and manage the NVMe subsystem as well as troubleshooting NVMe related issues.

What's NVMe?

NVM Express, NVMe, or Non-Volatile Memory Host Controller Interface Specification (NVMHCI), is a specification for accessing solid-state drives (SSDs) attached through the PCI Express (PCIe) bus. "NVM" stands as an acronym for non-volatile memory, which is used in SSDs. You can find detailed information on the support by Dell for Express Flash NVMe PCIe SSD and also information on use cases in the Dell Tech Center [blog](#).

In this paper we detail about the following main options that are available to monitor and manage the Express Flash NVMe PCIe SSD in VMware ESXi.

- Manage Express Flash NVMe PCIe SSD via CLI
- Manage Express Flash NVMe PCIe SSD via Webclient
- Manage Express Flash NVMe PCIe SSD Using Dell Openmanage Storage Services

NOTE: There is a pre-OS device management utility available as well in the form of HII (Human Interface Infrastructure) which is a standardized way of viewing and setting the device configuration. Refer to [express Flash NVMe PCIe SSD user guide](#).

Manage Express Flash NVMe PCIe SSD via CLI

This section list out few CLI options that may be useful in monitoring and managing Express Flash NVMe PCIe SSD in VMware ESXi. A white paper detailing management of PCIe SSD remotely can be found at the [Dell Tech Center](#).

NOTE: The same command lines may be used for the NVMe Express Flash PCIe SSD as well.

- **lspci | grep NVMe** command provides the details of the NVMe device connected to the system.

```
[root@static-10-10-201-221:~] lspci | grep NVMe
0000:08:00.0 Mass storage controller: Samsung Electronics Co Ltd Dell Express Flash NVMe XS1715 800GB PCIe SSD Controller [vmhba2]
[root@static-10-10-201-221:~] █
```

Figure 1 : **lspci | grep NVMe**

In this example, there is a 800GB Dell express flash NVMe PCIe SSD that is connected to the system.

- **esxcfg-scsidevs** command provides few useful options. It shows the device name associated for the NVMe device.

```
[root@static-10-10-201-221:~] esxcfg-scsidevs -a | grep NVMe
vmhba2 nvme link-r/a pscsi.vmhba2 (0000:08:00.0) Samsung Electronics Co Ltd Dell Express Flash NVMe XS1715 800GB PCIe SSD Controller
[root@static-10-10-201-221:~] █
```

Figure 2 : **esxcfg-scsidevs -a**

- **esxcfg-scsidevs -c**

This command displays the compact list for NVMe device such as Device UID, Device Type, Console Device, Size

```
[root@static-10-10-201-221:~] esxcfg-scsidevs -c -d t10.NVMe Dell_Express_Flash_NVMe_800GB S1J1NYAF40021000000001
Device UID Size Multipath Plugin Display Name Device Type Console Device
t10.NVMe Dell_Express_Flash_NVMe_800GB S1J1NYAF40021000000001 Direct-Access /vmfs/devices/disks/t10.NVMe Dell_Express_Flash_NVMe_800GB S1J1NYAF40021000000001
0000001 763097MB NMP Local NVMe Disk (t10.NVMe Dell_Express_Flash_NVMe_800GB S1J1NYAF40021000000001)
[root@static-10-10-201-221:~] █
```

Figure 3 : **esxcfg-scsidevs -c**

- **esxcfg-scsidevs -l -d**

This command shows the complete list of information for the NVMe device like firmware revision, Vendor, Model, Status, Is SSD etc. An example output is as below

```
[root@static-10-10-201-221:~] esxcfg-scsidevs -l -d t10.NVMe Dell_Express_Flash_NVMe_800GB S1J1NYAF40021000000001
t10.NVMe Dell_Express_Flash_NVMe_800GB S1J1NYAF40021000000001
Device Type: Direct-Access
Size: 763097 MB
Display Name: Local NVMe Disk (t10.NVMe Dell_Express_Flash_NVMe_800GB S1J1NYAF40021000000001)
Multipath Plugin: NMP
Console Device: /vmfs/devices/disks/t10.NVMe Dell_Express_Flash_NVMe_800GB S1J1NYAF40021000000001
Devfs Path: /vmfs/devices/disks/t10.NVMe Dell_Express_Flash_NVMe_800GB S1J1NYAF40021000000001
Vendor: NVMe Model: Dell Express Fla Revis: IPM0
SCSI Level: 6 Is Pseudo: false Status: on
Is RDM Capable: false Is Removable: false
Is Local: true Is SSD: true
Other Names:
vml.01000000002020202020202053314a314e59414634303032313044656c6c2045
VAAI Status: unknown
[root@static-10-10-201-221:~] █
```

Figure 4 : **esxcfg-scsidevs -l -d**

- **esxcfg-module -g nvme**

This command provides the loaded status of NVMe module (Driver) and the options set for the module. **nvme** is the driver name for all the types of NVMe PCIe SSD devices connected to VMware ESXi.

```
[root@static-10-10-201-221:~] esxcfg-module -g nvme
nvme enabled = 1 options = 'nvme_log_level=5'
[root@static-10-10-201-221:~] █
```

Figure 5 : **esxcfg-module -g nvme**

- **esxcfg-module -i nvme**

This command shows the NVMe module full information like Driver version, namespace and the various parameters for setting up the NVMe module options.

```

[root@static-10-10-201-221:~] esxcfg-module -i nvme
esxcfg-module module information
input file: /usr/lib/vmware/vmkmod/nvme
License: BSD
Version: 1.0e.0.35-2vmw.600.0.0.2159203
Name-space:
Required name-spaces:
  com.vmware.vmkapi@v2_3_0_0
Parameters:
  admin_cpl_queue_size: int
    NVMe number of Admin completion queue entries
  admin_sub_queue_size: int
    NVMe number of Admin submission queue entries.
  io_command_id_size: int
    Number of command IDs per submission queue
  io_cpl_queue_size: int
    NVMe number of IO completion queue entries
  io_sub_queue_size: int
    NVMe number of IO submission queue entries
  max_io_request: int
    IO Queue high water marker throttle
  max_namespaces: int
    Maximum number of namespaces supported.
  max_prp_list: int
    Maximum number of PRP lists supported.
  nvme_compl_worlds_num: int
    Total number of NVMe completion worlds/queues.
  nvme_dbg: int
    Driver NVME_DEBUG print level minimum 0 maximum 4
  nvme_force_intx: int
    Force using INTX mode instead of MSIx.
  nvme_log_level: int
    Log level.
      1 - error
      2 - warning
      3 - info (default)
      4 - verbose
      5 - debug
  transfer_size: int
    Maximum IO request Xfer size in 1K bytes maximum 1024
[root@static-10-10-201-221:~] █

```

Figure 6 : esxcfg-module -i nvme

- **esxcli storage core adapter stats get**

This command helps to view the granular details of how many writes and reads happened to the device and also provides the number of failed commands. This helps to troubleshoot nvme related issues.

```
[root@static-10-10-201-221:~] esxcli storage core adapter stats get -a vmhba2
vmhba2
  Successful Commands: 7491
  Blocks Read: 192
  Blocks Written: 0
  Read Operations: 177
  Write Operations: 0
  Reserve Operations: 0
  Reservation Conflicts: 0
  Failed Commands: 9
  Failed Blocks Read: 0
  Failed Blocks Written: 0
  Failed Read Operations: 0
  Failed Write Operations: 0
  Failed Reserve Operations: 0
  Total Splits: 0
  PAE Commands: 0
[root@static-10-10-201-221:~] █
```

Figure 7 : esxcli storage core adapter stats

- **esxcli storage core device list**

This command provides a set of attributes for the selected device.

```
[root@static-10-10-201-221:~] esxcli storage core device list -d t10.NVMe_____Dell_Express_Flash_NVMe_800GB_____S1J1NYAF40021000000001
t10.NVMe_____Dell_Express_Flash_NVMe_800GB_____S1J1NYAF40021000000001
  Display Name: Local NVMe Disk (t10.NVMe_____Dell_Express_Flash_NVMe_800GB_____S1J1NYAF40021000000001)
  Has Settable Display Name: true
  Size: 763097
  Device Type: Direct-Access
  Multipath Plugin: NMP
  Devfs Path: /vmfs/devices/disks/t10.NVMe_____Dell_Express_Flash_NVMe_800GB_____S1J1NYAF40021000000001
  Vendor: NVMe
  Model: Dell Express Fla
  Revision: IFW0
  SCSI Level: 6
  Is Pseudo: false
  Status: on
  Is RDM Capable: false
  Is Local: true
  Is Removable: false
  Is SSD: true
  Is WWOL PE: false
  Is Offline: false
  Is Perennially Reserved: false
  Queue Full Sample Size: 0
  Queue Full Threshold: 0
  Thin Provisioning Status: yes
  Attached Filters:
  VAAI Status: unknown
  Other UUIDs: vml.01000000000202020202053314a314e59414634303032313044656c2045
  Is Shared Clusterwide: false
  Is Local SAS Device: false
  Is SAS: false
  Is USB: false
  Is Boot USB Device: false
  Is Boot Device: false
  Device Max Queue Depth: 2048
  No of outstanding IOs with competing worlds: 32
  Drive Type: unknown
  RAID Level: unknown
  Number of Physical Drives: unknown
  Protection Enabled: false
  PI Activated: false
  PI Type: 0
  PI Protection Mask: NO PROTECTION
  Supported Guard Types: NO GUARD SUPPORT
  DIX Enabled: false
  DIX Guard Type: NO GUARD SUPPORT
  Emulated DIX/DIF Enabled: false
[root@static-10-10-201-221:~] █
```

Figure 8 : esxcli storage core device list

Manage Express Flash NVMe PCIe SSD via Webclient

Using webclient, NVMe devices can be managed in different ways. When an NVMe PCIe SSD is connected to ESXi, it is listed under the Storage Adapters and Storage Devices tab.

10.10.201.221

Actions

Getting Started

Summary

Monitor

Manage

Related Objects

Settings

Networking

Storage

Alarm Definitions

Tags

Permissions

Dell VSM

Storage Adapters

Storage Devices

Host Cache Configuration

Protocol Endpoints

Storage Devices

Filter

Name	Type	Capacity	Operational State	Hardware Acceleration	Drive Type	Transport
SEAGATE Serial Attached SCSI Disk (naa.5000c50057b932a7)	disk	3.64 TB	Attached	Unknown	HDD	Block Adapter
SEAGATE Serial Attached SCSI Disk (naa.5000c500102ee4a3)	disk	1.82 TB	Attached	Unknown	HDD	Block Adapter
Local USB Direct-Access (mpx.vmhba32:C0:T0:L1)	disk	0.00 B	Attached	Not supported	HDD	Block Adapter
DELL Serial Attached SCSI Disk (naa.6c81f660e230ad001bd3e5...	disk	278.88 GB	Attached	Not supported	HDD	Block Adapter
Local USB CD-ROM (mpx.vmhba32:C0:T0:L0)	cdrom		Attached	Not supported	HDD	Block Adapter
Local NVMe Disk (t10.NVMe____Dell_Express_Flash_NVMe_80...	disk	745.21 GB	Attached	Unknown	Flash	Parallel SCSI
DELL Serial Attached SCSI Disk (naa.6c81f660e230ad001bd034...	disk	3.64 TB	Attached	Not supported	HDD	Block Adapter
TOSHIBA Serial Attached SCSI Disk (naa.500003946c881250)	disk	186.31 GB	Attached	Unknown	Flash	Block Adapter
DELL Serial Attached SCSI Disk (naa.6c81f660e230ad00f00004...	disk	931.00 GB	Attached	Not supported	HDD	Block Adapter

Device Details

Properties

Paths

General

Name	Local NVMe Disk (t10.NVMe____Dell_Express_Flash_NVMe_800GB_____S1J1NYAF40021000000001)
Identifier	t10.NVMe____Dell_Express_Flash_NVMe_800GB_____S1J1NYAF40021000000001
Type	disk
Location	/vmfs/devices/disks/t10.NVMe____Dell_Express_Flash_NVMe_800GB_____S1J1NYAF40021000000001
Capacity	745.21 GB
Drive Type	Flash
Hardware Acceleration	Unknown
Transport	Parallel SCSI
Owner	NMP

Figure 9 : Properties of NVMe device(s)

The multiple usages of nvme device are explained in the following sections:

- *NVMe PCIe SSD as a VMFS datastore*

NVMe PCIe SSD can be formatted as a Virtual Machine File System (VMFS) datastore and the virtual machines can be placed on top of this for a better throughput.

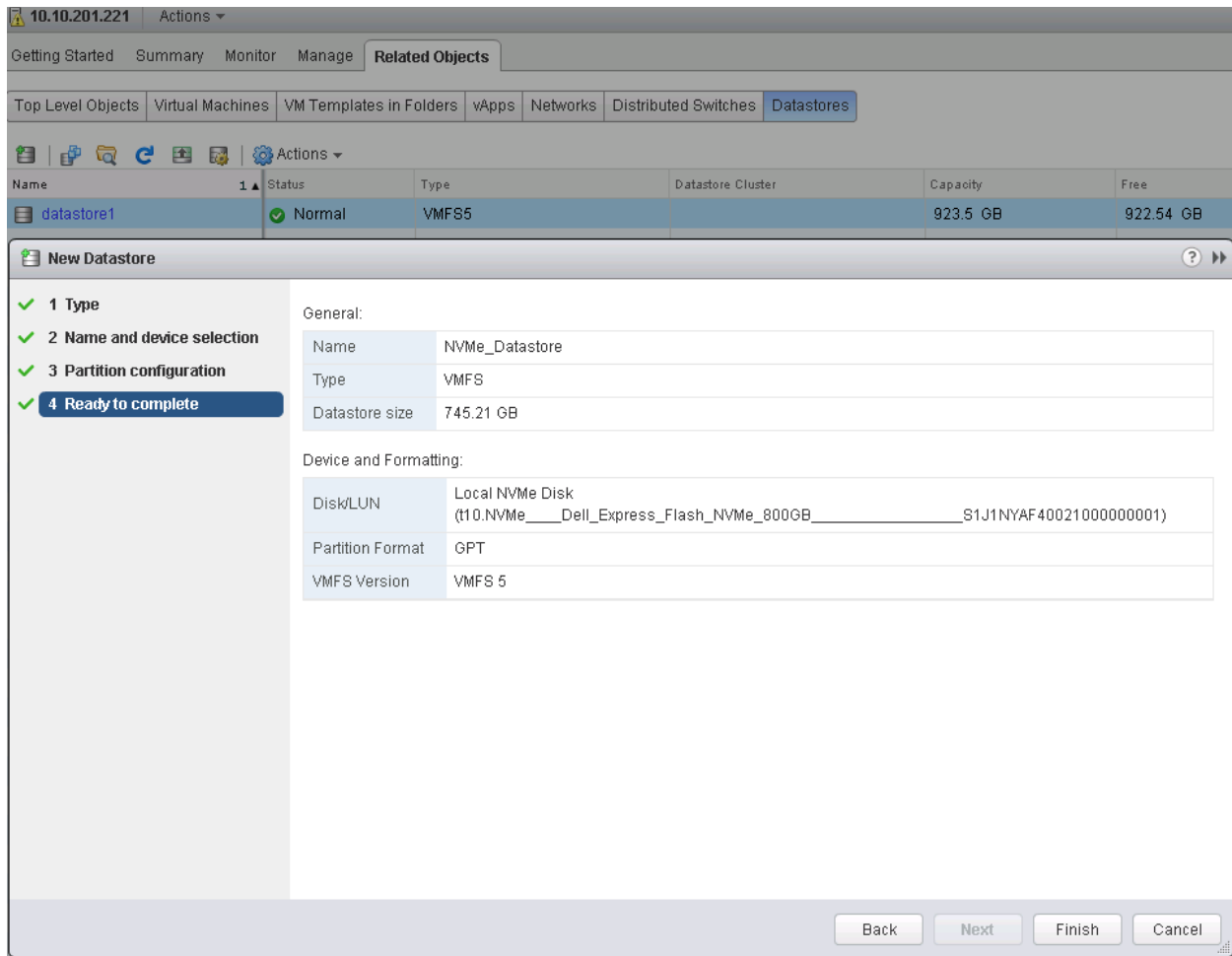


Figure 10 : Formatting NVMe device as a VMFS datastore

- *Pass-through of NVMe device(s) to the virtual machines*

It is possible to passthrough the NVMe device directly to the virtual machines running on ESXi. In this case, the guest operating system should have the NVMe driver installed and loaded to the kernel to make use of the device. Refer to VMware KB 1010789.

Note: This may not be a VMware supported configuration as VMware officially supports a limited number of devices for PCI passthrough feature.

- *NVMe PCIe SSD as Virtual Flash*

To manage the NVMe device as Vflash resource, refer to VMware KB 2059285. The screenshot displayed below is from a webclient which provides a look and feel of how the virtual flash resource allocation can be done.

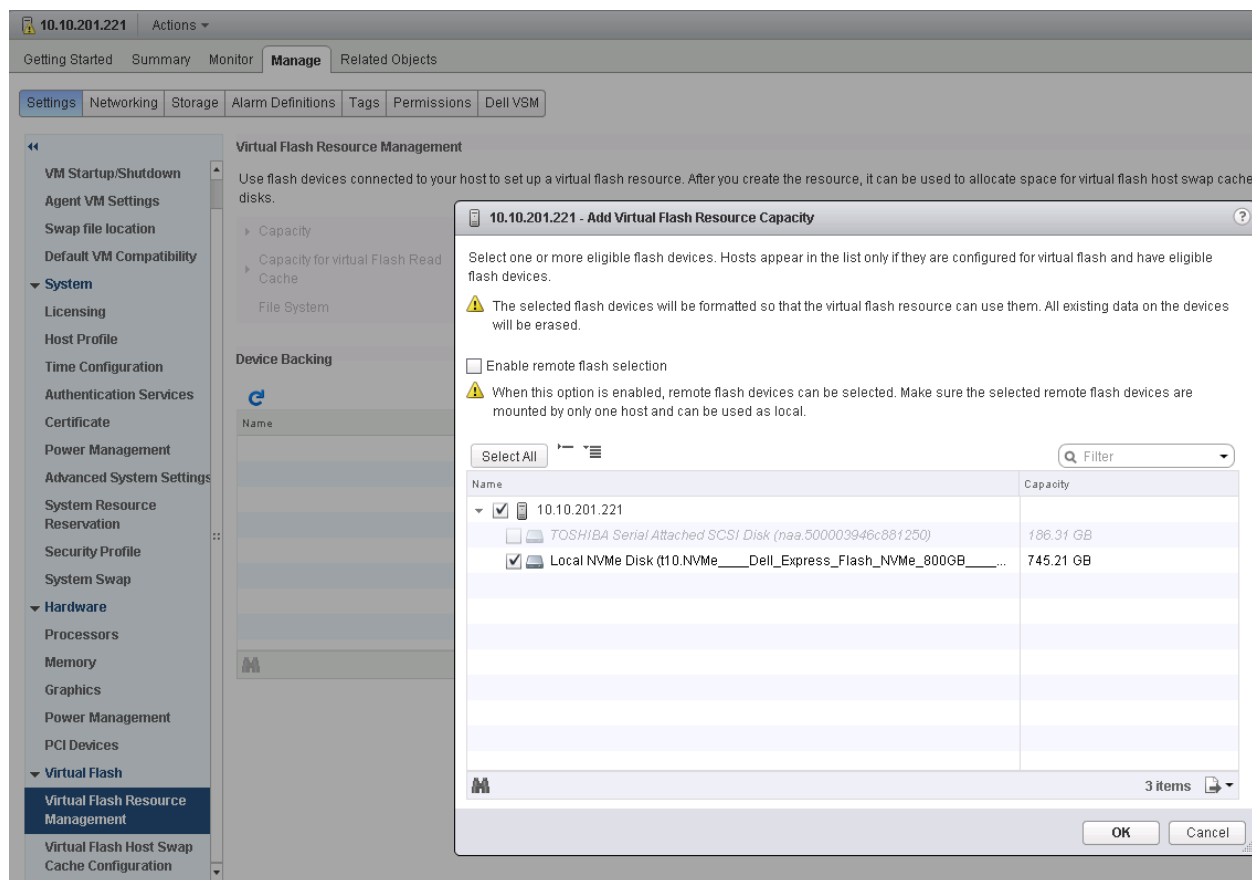


Figure 11 : NVMe device as a virtual flash

Manage Express Flash NVMe PCIe SSD using Dell Openmanage Storage Services

Dell [OpenManage version 8.1.0](#) release supports a range of servers such as R730, T630, R930,... which come with Express Flash NVMe PCIe SSD devices. This release added inband management for NVMe devices shipped with Dell servers on VMware ESXi 6.0.

The following form factors of the NVMe devices are supported by Dell OpenManage currently:

- 2.5 Inch disk
- Add-in card

2.5 Inch form factor devices are attached to the back plane of the Dell servers. Add-in cards are like Adapters and can be connected to PCI slots. OpenManage provides enumeration, configuration and monitoring of NVMe devices. GUI support and SNMP trap support are available from VMware ESXi 6.0 with respect to OpenManage perspective.

Significant Enumerated Properties of NVMe Devices:

- Device State
- Device Name
- Serial Number
- Remaining Rated Write Endurance

- Firmware Version
- Driver Version
- Capacity
- Model Number
- PCIe Negotiated Link Speed
- PCIe Maximum Link Speed
- Device Form Factor

The screen shot displayed illustrated the enumerated properties of NVMe devices.

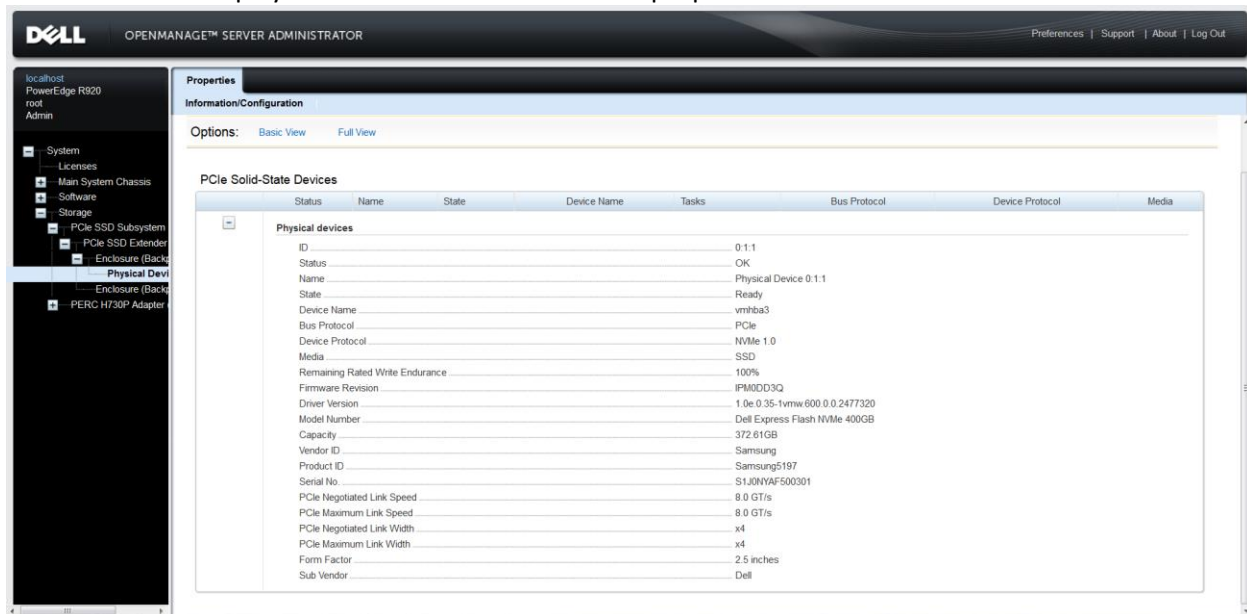


Figure 12 : Dell OpenManage UI displaying NVMe device(s)

Supported Operations:

- Cryptographic Erase
- Prepare To Remove
- Export Log
- Blink and Un-blink of devices

Blinking, Un-blinking and Prepare to Remove tasks are not available for Add-in card devices. Below Image shows the supported operations for NVMe devices of 2.5 Inch form factor.

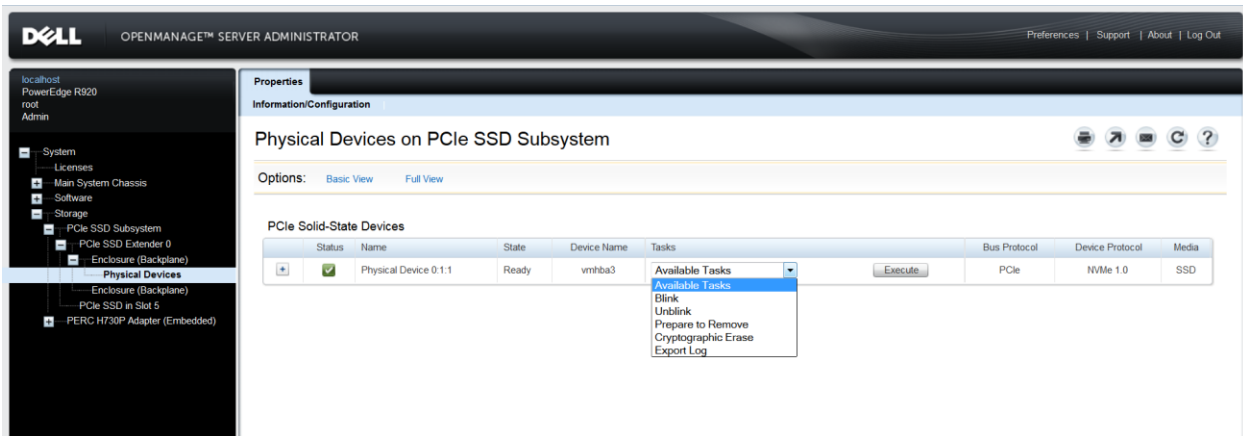


Figure 13 : OpenManage available tasks for NVMe device(s) of 2.5" form factor

Below screen shot shows the available operation for NVMe devices of Add-In card form factor.

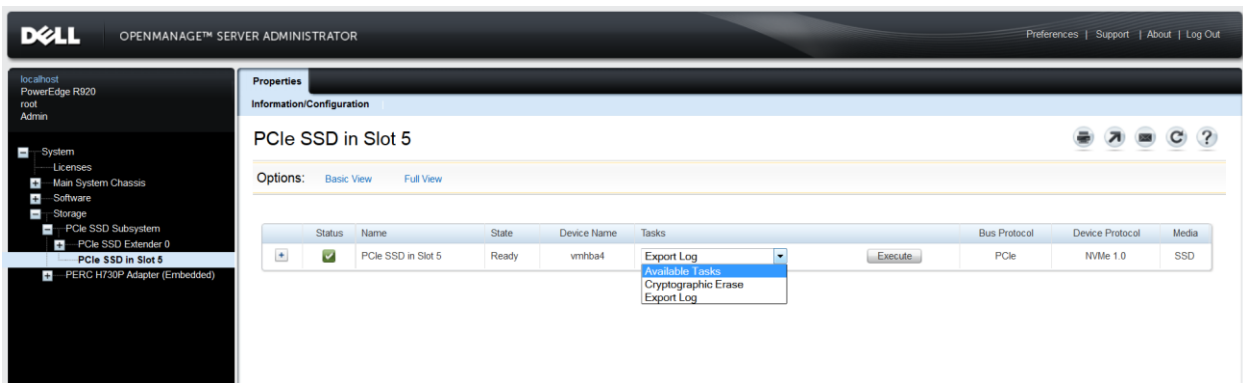


Figure 14 : OpenManage available tasks for NVMe device(s) of add-in card form factor

Monitoring:

OpenManage shows the devices status and state on GUI console. The state of the device is set to different values like READY, NOT-READY, FAILED, READ ONLY depending different properties/ SMART attributes of the NVMe devices.

OpenManage also generates alerts on console for NVMe devices when

- The device fails
- Critical temperature threshold of the device is exceeded
- Reliability of the device degrades
- The device is no longer functional
- On successful completion of Cryptographic erase Operation

OpenManage generates SNMP traps corresponding to these alerts which can be listened from other systems if configured.

Conclusion

This white paper provides an insight into the available utilities to monitor and manage Dell PowerEdge Express Flash NVMe PCIe SSD device(s) in VMware ESXi. The paper details few command line utilities which may be helpful while troubleshooting NVMe related issues followed by details around webclient and Dell OpenManage which may be used for extensive monitoring and managing.

References

- [Dell PowerEdge Express Flash NVMe PCIe SSD support in VMware ESXi](#)
- [Dell PowerEdge Express Flash NVMe PCIe SSD User's Guide](#)
- www.dell.com/poweredge/expressflash