Technical white paper

HP XP7 Storage and VMware vSphere 5 Best Practices and Configuration Guide

Table of contents

Executive summary

When supported with the correct underlying storage platform, server virtualization delivers increased infrastructure consolidation, administrative efficiency, business continuity, and cost savings. As a result, server virtualization is not only transforming the data center, but also the businesses that those data centers fuel. However, these transformative results depend on enterprise class storage to deliver the performance, availability, and flexibility to keep up with the dynamic and consolidated nature of virtualized server environments.

Storage is a critical component of the virtualization infrastructure. The storage platform must deliver high performance and use its capacity efficiently. Storage must also be modular and scalable to enable purchasing only what is needed, when it is needed. No single point of failure should exist, and failed components must be replaceable without interruption of service.

The latest generation of HP XP7 Storage is designed and built to exceed the economic and operational requirements of virtualized data centers by providing the storage performance, scalability, availability, and simplified management needed by enterprise businesses with growing storage needs. In this white paper, we explore the unique capabilities and integration of HP XP7 Storage with VMware vSphere. In addition, the white paper covers the best practice settings and configurations needed to optimize HP XP7 Storage for deployment with VMware vSphere. When deployed together, HP XP7 Storage with VMware vSphere provides enterprise businesses the benefits of increased consolidation savings by increasing virtual machine (VM) density, lowering storage costs, and realizing time savings from simplified storage management and provisioning.

As data center server and storage consolidation grows, the risk of production outages increases. Consolidation also increases the workload on servers and storage to new levels. Highly available and high-performance storage and server solutions are an integral part of the goal of an always available cloud-computing environment. With the HP XP7 Storage array (XP7), HP BladeSystem components, and VMware vSphere 5.5, a highly available and high-performance cloud infrastructure can be achieved.

This white paper outlines a virtualized infrastructure and offers best practices for planning and deploying the XP7 with VMware vSphere 5 on HP ProLiant BladeSystem servers. New VMware vSphere 5 features—such as Pluggable Storage Architecture, vStorage APIs for Array Integration (VAAI), VMware ESXI Server File System (VMFS) dynamic expansion, Virtual Machine Disk Format (VMDK) Thin Provisioning and Paravirtualized SCSI controllers—in combination with the XP7 dramatic performance improvements, Thin Provisioning (ThP), and flexibility help resolve any challenges consolidation and availability have to offer. This white paper serves as an additional resource aid for IT professionals responsible for implementing a VMware-virtualized environment. HP strongly recommends that you also thoroughly review documentation supplied with individual solution components to gain in-depth knowledge necessary for a comprehensive and reliable solution.

Target audience

This configuration guide is intended for IT Administrators and Solution Architects planning a virtualization deployment with HP XP7 Storage array. This and other documents pertaining to virtualization with HP are available here: hp.com/go/virtualization. More information regarding HP Storage solutions is available here: hp.com/go/storage/

Introduction

This document focuses on an enterprise server virtualization solution that offers benefits when compared to traditional environments. HP delivers a comprehensive portfolio of server, storage, and networking solutions, well suited for virtualization. This white paper focuses on the storage aspects of virtualization using HP XP7 Storage array. HP XP7 Storage is an enterprise-level storage solution that supports mixed application workloads, multi-tenancy requirements, converged protocols, and fibre channel (FC) storage networks.

Virtualization overview and challenges

Server virtualization is running multiple VMs on a single physical server where the hypervisor decouples resources—including memory, CPU, and network and storage devices—from the physical server and creates a virtual stack of hardware for the creation of VMs. Encapsulating the VM (operating system, application, data, and configuration settings) into a set of files enables mobility and allows multiple copies of the VM to be created for business resilience and disaster recovery (DR). Server virtualization ultimately enables dynamic flexibility while decreasing the economic and operational issues of infrastructure silos.

The resulting logical resource pool helps to reduce single points of failure, reducing vulnerabilities and improving resource use. However, to gain the benefits of advanced mobility functions, a networked storage environment must be created to facilitate automated and streamlined data management of storage that is linked to multiple places. You need to move data, not just compute resources, if you are to successfully move to another platform or location—for example, to sustain your business, whether for maintenance, or DR.

The point is to keep business going. In today's data centers, systems redundancy and business continuity is enabled by virtualization features (for example, VMware High-Availability, vMotion, vSphere Replication, and vCenter Site Recovery Manager in VMware) that work with and enhance the flexibility of the storage platforms. The data storage infrastructure is clearly an important building block in this scenario. As previously mentioned, VMware offers key high availability and business-continuity functionalities that rely on networked storage for mobility, high availability, and resource sharing. In order to achieve maximum scalability and peak resource utilization, it becomes necessary to network the storage infrastructure to keep pace with consolidation initiatives.

Main challenges to a successful virtualization strategy/solution

- Depleted resources—result in diminished performance and availability
- Lack of application awareness—OS virtualization doesn't virtualize the application
- Additional unexpected costs—cause the virtual solution to cost more than the physical problem
- Under-utilized virtualization features—the network limits implementation
- Increased storage requirements—VM growth exceeded planning
- Congested storage network
- Management complexity—management tools lack interoperability

HP XP7 Storage array—key features

- The HP XP7 Storage mitigates the risk of data loss and downtime because all components are redundant, hot swappable and firmware, and software can be upgraded online.
- The HP XP7 Storage supports a clustering solution that integrates remote mirroring with a high-availability server cluster to provide multisite server and storage DR.
- The HP XP7 Storage protects your data. Data held in cache is backed up to a Solid State Drive (SSD) to protect it from loss in the case of a power outage. RAID 6 adds fault tolerance at disk drive level.
- HP Continuous Track checks the HP XP7 Storage for unseen issues. The end-to-end checksum capability confirms data integrity from host port to disk and back.
- Security features in the HP XP7 Command View Advanced Edition software decrease your exposure to data loss. This software reduces the risk of unauthorized access, prevents unauthorized modification, and protects your information at the end of its lifecycle by completely deleting data on a specified volume.

Reduce cost and complexity of storage through efficient data management

The HP XP7 Storage system is a large enterprise-class storage system designed for organizations that simply cannot afford any downtime or tolerate any data loss. The XP7 reduces the risk of business downtime by providing a bulletproof platform with complete hardware redundancies, hot-swappable components, and non-disruptive online upgrades. Through thin provisioning, organizations can reduce costs by supplying storage capacity to applications from a virtualized pool. As a virtualization platform, the XP7 enables important functions such as heterogeneous data migration, array repurposing, and storage tiers. The XP7 also eliminates contention for valuable storage resources through array partitioning. Partitioning enhances the success of storage consolidation efforts by assigning array resources to certain applications and administrators.

- Increase data center efficiency using virtualization with thin provisioning, heterogeneous external storage, local and remote multisite replication, automated smart-tiered capacity, and online data migration.
- Simplify management with an easy-to-use, task-based GUI; common and consistent CLI; easy serviceability, trouble-free provisioning with one-click volume creation, and dynamic or automatic provisioning and optimization.
- Improve use of SSD and Flash storage media and automate data movement between storage tiers with smart tiers software, saving administration time, reducing complexity, increasing performance, and maximizing storage media investments.
- Save power and cooling costs with thin provisioning and smart tiers that decrease stranded storage space.

VMware VAAI storage performance benefits

The vSphere Storage APIs are a set of technologies and interfaces that enable vSphere to leverage storage resources to deliver improved efficiency, control, and ease of use. The vSphere Storage APIs for Array Integration (VAAI) is one of these technologies. The VAAI initiative introduces APIs to improve performance, resource usage, and scalability by leveraging more efficient storage array-based operations.

VAAI helps reduce the storage bandwidth consumed by a vSphere host and improves data center scalability. Storage operations like VM provisioning, Storage vMotion, and virtual disks creation, will consume less CPU, memory, and network bandwidth when using the VMware VAAI-compliant HP XP7 Storage system. With hardware offload, vSphere hosts may perform certain operations faster while consuming fewer host resources.

Storage hardware primitives for VAAI

Primitives are specific functions used with VAAI that serve as integration points to storage arrays. VMware develops the primitives and array vendors integrate the primitives into their arrays. Primitives in VAAI allow the hypervisor to communicate directly with storage arrays. Their purpose is to allow close integration between the hypervisor and storage array as well as to offload storage functionality traditionally handled by the hypervisor. Storage arrays can handle these functions more intelligently and efficiently because they are purpose built to perform storage tasks and can complete the request much faster than the host.

Tip

To confirm that VAAI is supported, a number of `ESXi cli` commands can be executed from the vSphere host (SSH session). For example, to see which primitives are supported for a particular device, use the following command.

```
# esxcli storage core device vaai status get -d naa.xxx
```

For more information on VMware vStorage APIs for Array Integration (VAAI), refer to the VMware Knowledgebase article FAQs for VAAI and the VMware Community VAAI document.

Test environment

The performance and best practices described in this document are derived from a test environment comprised of HP ProLiant c-Class blade servers with Virtual Connect, HP ProLiant DL-class rackmount servers, an HP XP7 Storage system, and VMware vSphere 5.5. The solution configuration allows for performance tuning by using multiple paths, queue depth adjustments, vSphere features such as Paravirtualized SCSI adapters, and more. This solution also provides an environment that can be used to understand capacity considerations such as Thin Provisioning and volume growth.

Table 1. Equipment details of vSphere/HP XP7 Storage test environment

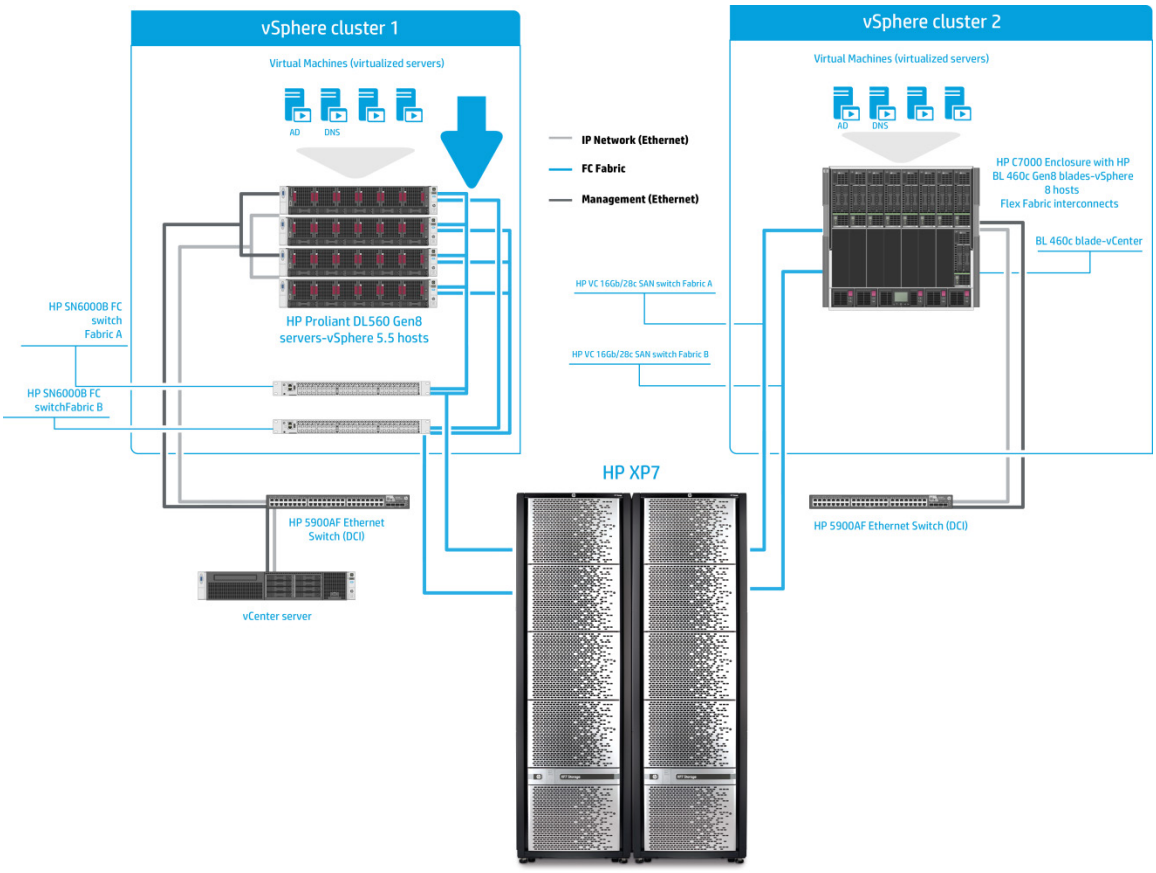
	Model	**Configuration**
vSphere host server	HP ProLiant DL560 Gen8	4 Xeon 8-core processors
Host Bus Adapter (HBA)	Emulex LPe 12000	8 GB FC
Storage provisioned	HP XP7 with 10k RPM SAS drives	RAID5 (7D+1P) x 5.5 TB parity groups
Storage connectivity/fabric	HP SN6000B 16Gb FC switch	X 2 (two fabrics)
Storage connectivity	HP XP7/4 HBAs per power CHA	4 FC ports per controller
vSphere host server	HP ProLiant BL460c Gen8	2 Xeon 8-core processors
HBA	QLogic QMH2672 16GB mezzanine cards	16 GB FC
Storage provisioned	HP XP7 with 10k RPM SAS drives	RAID5 (7D+1P) x 5.5 TB parity groups
Storage connectivity	HP Virtual Connect 16GB SAN switches (4)	2 FC ports per SAN switch connected to XP7 array
Network connectivity	HP 5900AF-48XGT 48-port Ethernet switch (JG336A)	48 RJ-451/10GbE ports 4 QSFP+ GbE ports

Information regarding HP storage networking (FC switches) can be found here: hp.com/go/storefabric

For more information regarding HP network switches for the data center, visit: hp.com/go/networking

More information regarding HP ProLiant servers is available here: hp.com/go/proliant

Figure 1. The test environment topology with two vSphere 5 clusters and the HP XP7 Storage array



VMware environment

- Servers: Four HP ProLiant DL560 Gen8 servers configured as vSphere 5.5 hosts in a vSphere cluster. Connectivity to storage was via Emulex LPE12000 Fiber-Channel host bus adapters (HBA).
- Blade servers: Eight HP ProLiant DL460c Gen8 servers configured as vSphere 5.5 hosts in a vSphere cluster. Connectivity to storage was via Qlogic QMH2672 Fiber-Channel HBAs.
- Management: VMware vCenter Server installed on an HP ProLiant DL360 Gen8 server. Blade environment managed by vCenter Server installed on an HP ProLiant BL460c Gen8 blade.

Important

Hardware drivers supplied with the HP-supplied vSphere/ESXi build include the necessary HBA (host bus adapter) drivers for storage connectivity. See this HP website for more details on the drivers included.

To use the HP-supplied vSphere 5.5 image, see the HP ProLiant website. The new **ProLiant Gen8 servers** require the use of the HP Customized image for a successful installation. The drivers for the new network and storage controllers in the Gen8 servers are integrated in the HP customized image and are not part of the generic ESXi image that is distributed by VMware. ESXi requires drivers for these controllers to be integrated because you are not able to insert them during installation. Compatible HP bundles are available for VMware Image Builder, Update Manager, and ESXi CLI.

Composition of virtual machines

Three types of VMs listed below were configured and built for our testing exercises.

- Windows® 2012 Server (100GB eager-zeroed thick or EZT VMDK, 80GB EZT VMDK, and 30GB thin-provisioned VMDK)
- Windows 2008 R2 Server (100GB EZT VMDK, 80GB EZT VMDK, and 25GB thin-provisioned VMDK)
- Ubuntu Server version 12 (20GB VMDK)

The Windows Server VMs were configured with both thin-provisioned virtual disks (VMDK) and eager-zeroed thick (EZT) VMDK of various sizes. This Windows VM was also made into a template and used for cloning and vMotion test exercises. The Ubuntu Linux VMs were configured with an eager-zeroed thick 25GB VMDK with the Ubuntu Server version 12 operating system installed. A number of services such as Samba were also installed.

Installation roadmap

Follow these steps to install and configure the XP7 for a VMware vSphere environment:

1. Install and configure the disk array
   - Set the host mode and host group mode for the disk array ports
   - Set the System Option Modes
   - Configure the FC ports
   - Create Host Groups
2. Install and configure host servers (vSphere)
   - Load the operating system and software
   - Install and configure the HBAs
   - Configure clustering and fabric zoning
   - Configure fabric zoning and LUN security for the operating systems
   - Configure VMware vSphere
3. Connect the XP7
   - Connect FC cables to CHAs and FC switches
   - Define the paths
4. Set up VMs and guest operating systems
   - Set the SCSI disk timeout value for Windows VMs
   - Share logical units (LUNs)
   - Select the SCSI emulation driver

For more information about installation steps, refer to the *HP XP7 Provisioning for Open Systems Guide* and the *HP XP7 Remote Web Console User Guide*. These and other HP documents can be found on the following HP web sites:
hp.com/go/xp7

h20566.www2.hp.com/portal/site/hpsc

Configuration details

Determining which hardware devices and software and firmware versions are compatible can be a challenge. The HP Single Point of Connectivity Knowledge (SPOCK) website can help. SPOCK is a portal website dedicated to being a one-stop resource for SAN interoperability knowledge. For more information, go to the SPOCK website at hp.com/storage/SPOCK.

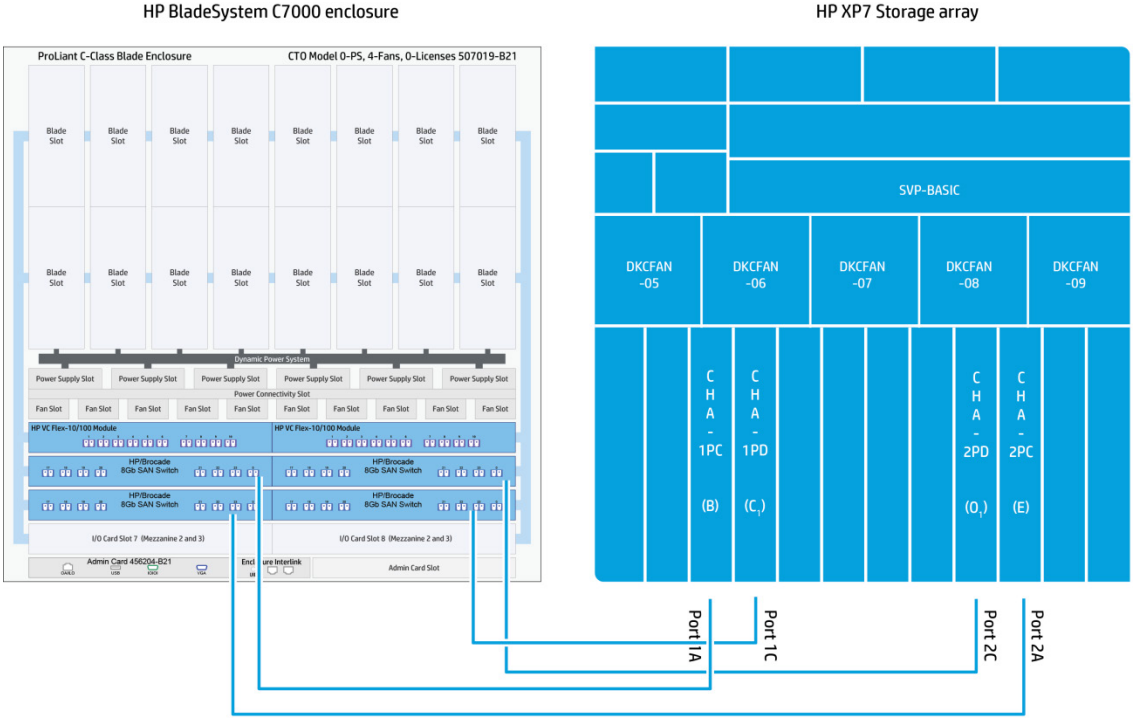
Fibre Channel connectivity

The vSphere host blade servers are each configured with a QLogic QMH2672 16Gb FC HBA. The QMH2672 is a dual-port adapter connected to a pair of Virtual Connect 16 Gb FC modules through the BladeSystem enclosure backplane. The result is two different fabrics. A configuration option is available to use four Virtual Connect Fibre-Channel modules, which would provide four different fabrics. The HP ProLiant DL560 servers installed as vSphere hosts in the other vSphere environment are each configured with Emulex LightPulse LPe12000 8 Gb FC HBA adapters, which are connected to a pair of HP SN6000B 16 Gb Fibre-Channel switches. Each SN6000B FC switch is connected via two FC cables to the HP XP7 Storage array.

The XP7 Storage array is configured to use four FC ports for our BladeSystem test environment. CHAs come in pairs of two blades; each pair resides in a separate power domain or cluster. Two ports on the Cluster 1 CHA blade of the array and two ports on the Cluster 2 CHA blade are used for the primary storage paths to the HP Virtual Connect FC switches in bay 3 and bay 5 of the HP BladeSystem enclosure (see figure 2). Additional FC ports were used to connect to the SN6000B FC switches in the other vSphere environment, consisting of the HP ProLiant DL servers (see figure 1). CHA blades contain microprocessor (MP) modules. An MP module contains two microprocessors and controls four ports (two ports per MP). To saturate the CHA microprocessors for experimental purposes, both ports on the front and rear CHA blades are configured to share a single MP. In production environments with high I/O demand, it is beneficial to configure the host ports to use separate microprocessors to enable higher performance.

Each CHA port is configured to use LUN Security with the Remote Web Console (RWC). LUN Security requires the creation of host groups containing the vSphere host servers. When sharing LUNs with multiple ESXi hosts in support of features such as VMotion, a storage administrator saves time by adding multiple vSphere hosts into one host group. An LDEV can then be configured into the vSphere host group at one time as opposed to one time per vSphere host. It is recommended to create a separate host group for each vSphere cluster; this provides a level of LUN security. In the test environment presented here, a host group was created which contains the WWNs for the blade servers, and another host group was created which contains the WWNs for the ProLiant DL560 rack-mount servers.

Figure 2. The FC connections from HP BladeSystem C7000 enclosure to the HP XP7 Storage system



A best practice for FC cabling for high-availability is to connect one host port on a channel host adapter (CHA) () from each power cluster/domain on the XP7 array to a port on the FC switch. This design reduces single points of failure in the SAN fabric configuration. In the example above, ports 1A and 2A are connected to the SAN switches in bays 3 and 5 of our HP BladeSystem C7000 enclosure. There are multiple configuration options for FC cabling an environment with the HP BladeSystem enclosure and SAN interconnect switches, refer to the HP Virtual Connect Fibre-Channel cookbook for more information.

Two ports on the Cluster 1 CHA blade of the array and two ports on the Cluster 2 CHA blade are used. In our example, ports 1A and 2A make a port pair, and are cabled into both fabric A (SAN switch 3 in bay 3) and fabric B (switch 5 in bay 5) as shown in figure 2.

Fabric zoning

The HP XP7 Storage system in our test environment has four XP7 8 Gb FC CHA ports connected to the SAN switches, providing for up to four paths per host. This configuration allows for enabling up to four different fabrics.

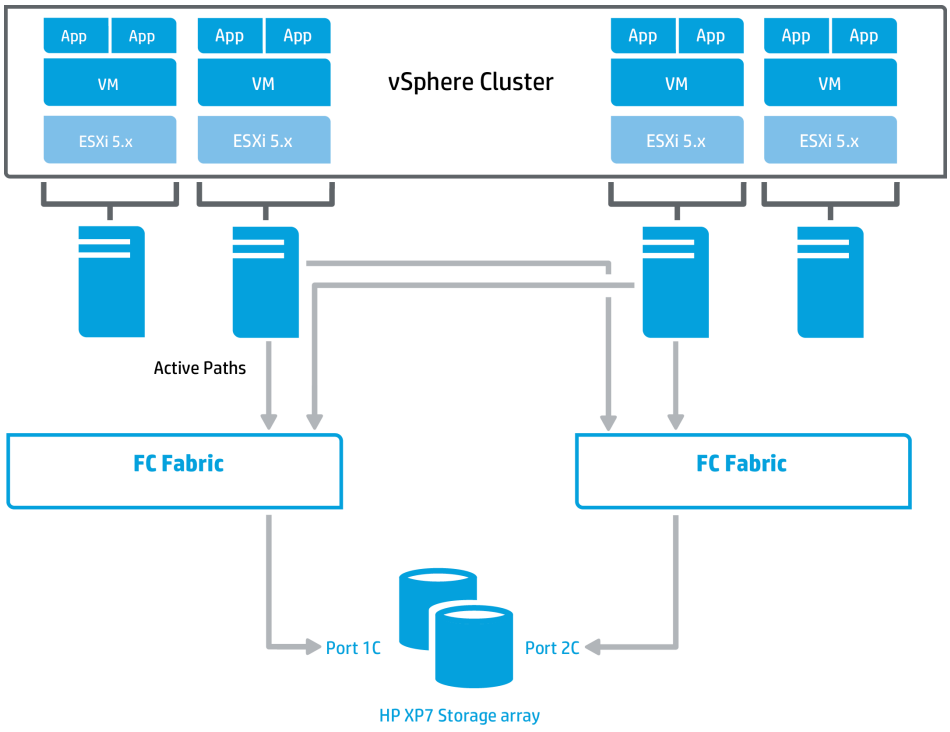
The zoning configuration is host-centric, with one host zone per fabric consisting of the host and its XP7 targets. Use a good zoning practice to reduce troubleshooting time, prevent resource conflict, and improve bus scan times. Because the host FC connection to the SAN switch is aggregated through the Virtual Connect module, N-port ID Virtualization (NPIV) is used. From the SAN switch view, it appears as one port with many worldwide names (WWN).

The WWN zoning and aliases used in our example configuration simplify management and future maintenance. An alias was created for each vSphere host/blade and a separate alias was created for the XP7 array, then the WWN of each HBA for the respective host was added to the alias. A zone was created and the host aliases and array alias were added to the zone. This was done for each FC switch in each fabric.

You can connect multiple clusters of various operating systems to the same switch and fabric using appropriate zoning and LUN security as follows:

- Storage port zones may overlap if more than one operating system needs to share an array port.
- Heterogeneous operating systems may share an XP array port if you use Secure Manager and set the appropriate host group and mode. All others must connect to a dedicated XP array port.
- Use Secure Manager for LUN isolation when multiple hosts connect through a shared array port. Secure Manager provides LUN security by allowing you to restrict which LUNs each host can access.

Figure 3. The HP XP7 and FC fabric zoning example

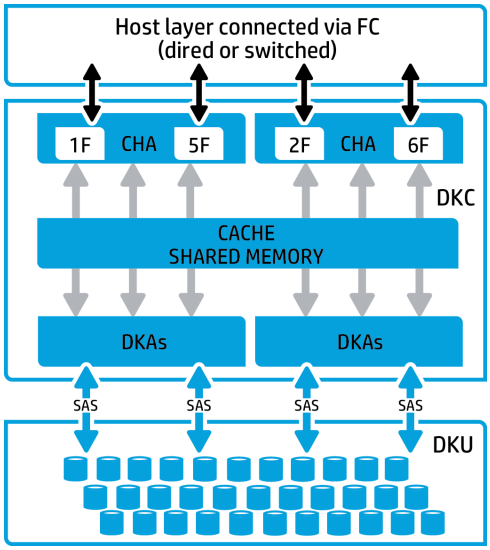


For more information about zoning and SAN design considerations, see the HP SAN design reference guide located here: HP SPOCK website (Single Point of Connectivity Knowledge for HP Storage).

HP XP7 internal storage paths

Connectivity from the back-end disks to the front-end host ports occurs through a series of controllers and cache. The physical disks are installed in groups of four (called an Array Group) into their disk cabinets, or DKU. Each disk connects to disk adapters (DKA) through Serial-Attached SCSI (SAS) , where processing occurs for I/O that comes from hosts using the XP7. Hosts connect through channel adapters (CHA) which help to manage interaction with the cache between them and the DKAs (see Figure 4).

Figure 4. HP XP7 Storage array I/O path



Array Groups are then configured into Parity Groups (also called RAID Groups) that can be configured as RAID 1, RAID 5, or RAID 6. RAID 1 configurations can be created in a 2D+2D (two data disks mirrored with two data disks) or 4D+4D Parity Group. RAID 5 Parity Groups offer the most capacity and performance flexibility, with four-disk (3D+1P or three data disks and one parity disk) configurations and eight-disk (7D+1P) configurations that can also be interleaved for up to 32 disks (28D+4P). RAID 6 Parity Groups can be created with eight disks in a 6D+2P configuration.

After the Array Groups are added to Parity Groups, the storage can then be divided into a presentable unit called a logical device (LDEV). LDEVs can be presented to hosts as LUNs and can be added to a Thin Provisioning pool.

HP XP7 host modes

Host mode sets a host group to work on the SAN in a method that is compatible with the OS using the provisioned devices. For VMware VMFS and raw device mapping (RDM) volumes, a host mode of 0x01 is used. Finally, if there are plans to use the XP7 ThP Online LUN Expansion capabilities, host mode option **40** must be set. For more information about host modes and OS support, go to the SPOCK website at hp.com/storage/SPOCK.

After the XP7 LDEV is presented to the vSphere host, run an HBA rescan to probe the bus for any new devices. When the device is discovered, it is entered as a record in the details for Storage Adapters in the vSphere Client. After new entries are loaded into the storage-adapter device details, helpful information about LUN numbers, capacity, and multipathing owner is available in the following command list. For ease of storage management, the LUN ID on the LDEV mapped to the vSphere hosts can be set to a specific ID number, making identification of the new LUN from the vSphere host much easier. See Presenting storage volume to vSphere hosts in this white paper. A more detailed look at the Network Address Authority (NAA) Identifier and Runtime Name reveals valuable information.

VMware vSphere 5.x obtains the 128-bit NAA identifier (disk UUID) from the Extended Vital Product Data (EVPD) located in the LUN INQUIRY page 0x83. The XP7 NAA identifier provides the array serial number, control-unit (CU) number, and LDEV number. In previous versions of VMware vSphere/ESXi, target shifting could cause the identifier of the LUN to change and cause persistency issues. The NAA identifiers use a unique name which allows a device to be discovered as a distinctive entry every time, regardless of a shifting target enumerator. See the section Create host groups and LUN mapping for more information regarding NAA identifiers. Administrators can access many such interfaces, commands, and files in vSphere that provide more information about their storage devices. Some common methods used in this project are listed here. A number of esxcli commands are listed (see Tip for esxcli).

- vSphere Client–open the Configuration Tab–under Storage Adapters, Storage, Manage Paths
- vSphere Client–Edit settings on the VM, and look at Hard Drives

```
# esxcfg-mpath
# esxcfg-scsidevs
# esxcli
# ls –l /vmfs/devices/disks
# esxcfg-scsidevs –u / -l / -m / -c
# cat /proc/scsi/scsi
# cat /proc/scsi/qla2xxx
```

When performing troubleshooting steps on storage, the following VMware documentation may be helpful.

blogs.vmware.com/vsphere/2012/05/troubleshooting-storage-performance-in-vsphere-part-1-the-basics-.html

blogs.vmware.com/vsphere/2012/06/troubleshooting-storage-performance-in-vsphere-part-2.html

blogs.vmware.com/vsphere/2012/06/troubleshooting-storage-performance-in-vsphere-part-3-ssd-performance.html

blogs.vmware.com/vsphere/2012/07/troubleshooting-storage-performance-in-vsphere-part-5-storage-queues.html

Figure 5. The vSphere host via SSH session and example of the esxcfg-mpath command above

```
~ # esxcfg-mpath --list-paths
naa.60060e8007273e000030273e00000012 : HP Fibre Channel Disk (naa.60060e8007273e000030273e00000012)
   vmhba1:C0:T0:L4 LUN:4 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:ad:29 WWPN: 50:01:43:80:24:d0:ad:
28  Target: WWNN: 50:06:0e:80:07:27:3e:00 WWPN: 50:06:0e:80:07:27:3e:00
   vmhba2:C0:T0:L4 LUN:4 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:ad:2b WWPN: 50:01:43:80:24:d0:ad:
2a  Target: WWNN: 50:06:0e:80:07:27:3e:42 WWPN: 50:06:0e:80:07:27:3e:42
   vmhba3:C0:T0:L4 LUN:4 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:b1:99 WWPN: 50:01:43:80:24:d0:b1:
98  Target: WWNN: 50:06:0e:80:07:27:3e:10 WWPN: 50:06:0e:80:07:27:3e:10
   vmhba4:C0:T0:L4 LUN:4 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:b1:9b WWPN: 50:01:43:80:24:d0:b1:
9a  Target: WWNN: 50:06:0e:80:07:27:3e:52 WWPN: 50:06:0e:80:07:27:3e:52

naa.600508b1001c4d819c649a4d1c95ecba : HP Serial Attached SCSI Disk (naa.600508b1001c4d819c649a4d1c95ecba)
   vmhba0:C0:T0:L1 LUN:1 state:active sas Adapter: 5001438022dd4cb0  Target: 1438022dd4cb0

naa.60060e8007273e000030273e00000001 : HP Fibre Channel Disk (naa.60060e8007273e000030273e00000001)
   vmhba1:C0:T0:L9 LUN:9 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:ad:29 WWPN: 50:01:43:80:24:d0:ad:
28  Target: WWNN: 50:06:0e:80:07:27:3e:00 WWPN: 50:06:0e:80:07:27:3e:00
   vmhba2:C0:T0:L9 LUN:9 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:ad:2b WWPN: 50:01:43:80:24:d0:ad:
2a  Target: WWNN: 50:06:0e:80:07:27:3e:42 WWPN: 50:06:0e:80:07:27:3e:42
   vmhba3:C0:T0:L9 LUN:9 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:b1:99 WWPN: 50:01:43:80:24:d0:b1:
98  Target: WWNN: 50:06:0e:80:07:27:3e:10 WWPN: 50:06:0e:80:07:27:3e:10
   vmhba3:C0:T1:L9 LUN:9 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:b1:99 WWPN: 50:01:43:80:24:d0:b1:
98  Target: WWNN: 50:06:0e:80:07:27:3e:12 WWPN: 50:06:0e:80:07:27:3e:12
   vmhba4:C0:T0:L9 LUN:9 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:b1:9b WWPN: 50:01:43:80:24:d0:b1:
9a  Target: WWNN: 50:06:0e:80:07:27:3e:52 WWPN: 50:06:0e:80:07:27:3e:52

naa.60060e8007273e000030273e00000003 : HP Fibre Channel Disk (naa.60060e8007273e000030273e00000003)
   vmhba1:C0:T0:L1 LUN:1 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:ad:29 WWPN: 50:01:43:80:24:d0:ad:
28  Target: WWNN: 50:06:0e:80:07:27:3e:00 WWPN: 50:06:0e:80:07:27:3e:00
   vmhba2:C0:T0:L1 LUN:1 state:active fc Adapter: WWNN: 50:01:43:80:24:d0:ad:2b WWPN: 50:01:43:80:24:d0:ad:
2a  Target: WWNN: 50:06:0e:80:07:27:3e:42 WWPN: 50:06:0e:80:07:27:3e:42
```

Tip

For more information on `esxcli` commands and their related options, see Getting Started with vSphere Command-Line Interfaces or the reference vSphere Command-Line Interface Concepts and Examples.

Prerequisites for esxcli:

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. For troubleshooting, run `esxcli` commands in the ESXi Shell or enable SSH on the vSphere host server, under Troubleshooting Options. See VMware Knowledgebase article 1017910 for information on enabling SSH or Tech Support mode on a vSphere 5.x host.

Multipathing and vSphere

Multiple, redundant paths from the vSphere host servers to the storage via the SAN are a key ingredient in a successful virtualization environment. Highly available and high-performance storage and server solutions are critical components in a vSphere configuration.

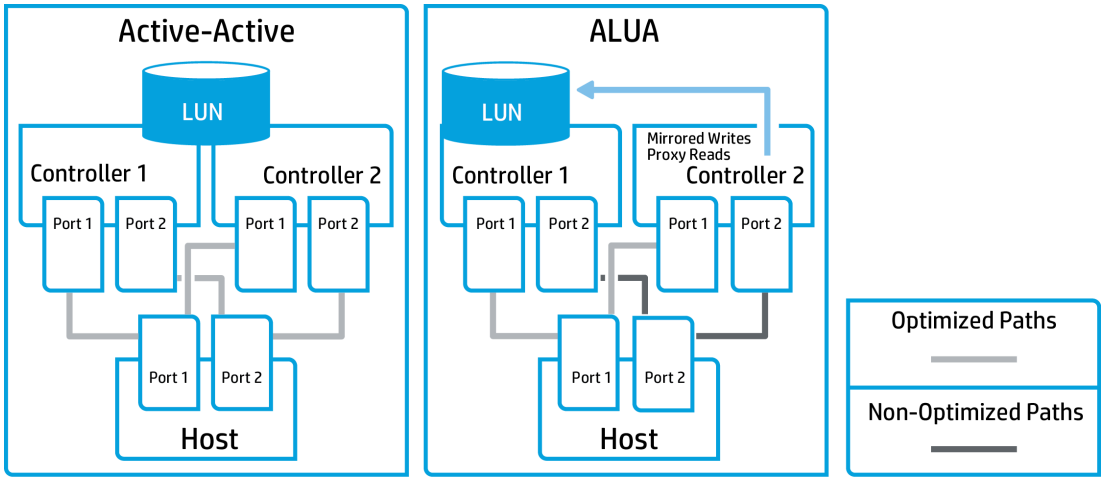
The latest storage enhancements in VMware vSphere 5.5 Pluggable Storage Architecture (PSA) carry with it a much improved multipathing capability that can save storage administrators time while improving performance and availability. In previous versions of VMware vSphere, it was possible to manage paths to LUNs from an array and provide some high availability, should a path fail; however leveraging a load-balancing policy was at best experimental. The lack of load-balancing policies requires a storage administrator to manage path use manually, making sure paths are not becoming overburdened and are evenly spread across all allocated disk array host ports. With vSphere 5, there is now support for round-robin load balancing across multiple paths, which not only provides excellent performance and availability, but can also provide administrative ease of use. You may no longer need to manually set a balanced load between the vSphere 5.5 host and its targets. Refer to figure 4 for an example of multiple paths from the vSphere hosts to the HP XP7 storage array.

Active-Active and ALUA

Before covering the PSA, it is important to identify the type of LUN availability the XP7 provides. There are some arrays that offer best performance by targeting the managing controller of a given LUN, or Asymmetric Logical Unit Access (ALUA); however, the XP7 is not an ALUA array. The XP7 offers best performance to a LUN through any XP host port that is accessible by the server, as shown in Figure 6.

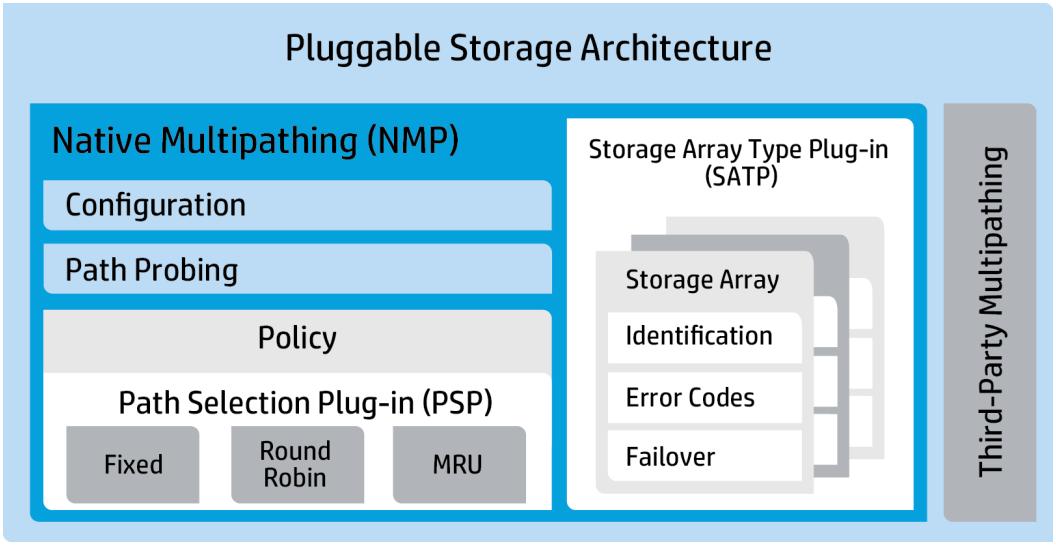
It is important to take note of the Active-Active LUN access for the XP7, because it is used in selecting the proper Storage Array Type Plug-in (SATP) for the array and because administrators can spend less time worrying about which ports are optimized for I/O intensive workloads.

Figure 6. Active-Active storage array connectivity compared to ALUA-storage array connectivity.



VMware's PSA is a modular framework enabling a base set of system multipathing policies or Native Multipathing (NMP) and third-party plug-ins that can be developed by a vendor. NMP in vSphere 5 works together with the SATP and the Path Selection Plug-in (PSP) to enable static path and load-balanced multipath solutions (see Figures 7 and 8). VMware PSPs handle the following operations: determine on which physical path to issue the I/O requests sent to a storage device; determine which path to activate next if the current physical path to the storage device fails.

Figure 7. VMware PSA framework with storage array integration



Native Multipathing
NMP handles many non-array specific operations such as periodical path probing, monitoring, and building the multipathing configuration. NMP is the common link between SATP and PSP.

Storage Array Type Plug-in
Storage Array Type Plug-in (SATP) is a plug-in that controls device discovery, processes array return codes when errors happen, and manages path failover. Because there is no HP-specific SATP for the XP7, a default SATP has been tested and is used.

NMP communicates with PSPs for the following operations:

- Set up a new logical storage device and claim the physical paths.
- Retrieve a set of active physical paths to be used for path selection.
- Select a physical path on which to issue the I/O requests for a specific storage device.
- Select a physical path to active when a path failure occurs.

The XP7 arrays use the SATP VMW_SATP_DEFAULT_AA, which can be seen in the vSphere client or in the ESXi console by performing the following command line operation:

```
# esxcli nmp satp list
```

Name	Default PSP	PSP Description
VMW_SATP_DEFAULT_AA	VMW_PSP_FIXED	Supports non-specific active/active…

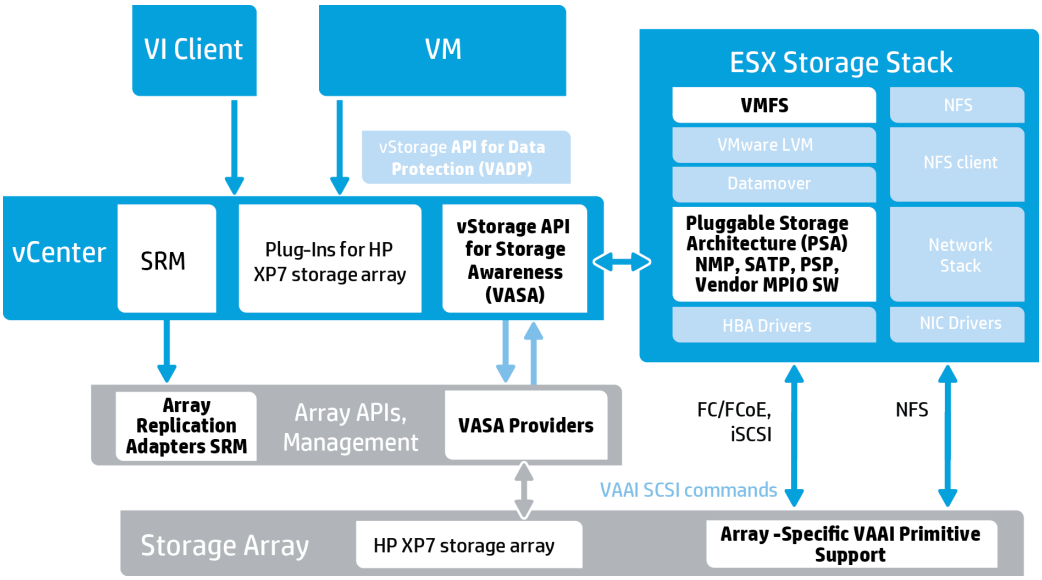
To see the mapping between the disk array and the SATP, issue the following command:

```
# esxcli nmp satp listrules
```

Results from this command will show the SATP and other information about vendor ID, model numbers, ALUA settings, and more.

The SATP shows that any new LUN for the XP7 will have a multipathing policy of FIXED, which is denoted by the PSP VMW_PSP_FIXED. It is important to note that the SATP settings affect all arrays on the VMware vSphere host that match as the VMW_SATP_DEFAULT_AA type. Although the SATP is a global setting on the host, the PSP can be used to modify the multipathing behavior for each LUN.

Figure 8. Integration points for PSA and vStorage APIs (VAAI) in the vSphere storage stack



Path Selection Plug-in (PSP)

The Path Selection Plug-in (PSP) handles the selection of the best available path to use for I/O requests and comes with the following three selections:

- Most Recently Used (MRU)—Uses the first available path discovered at boot time unless it is configured manually. If the preferred path is not available, an alternative available path is selected. The host will remain on the last set path. Think of this selection as a failover, without a failback.
- Fixed—Uses the first available path discovered at boot time unless it is configured manually. If the preferred path is not available, the next available path will be used until the preferred path becomes available, at which time it will fail back to the preferred path. Think of this selection as a failover-failback path.
- Round-robin—Uses a pre-determined path selection algorithm to manage I/O for all available paths.

To list the PSPs, from an SSH session to the vSphere host, or in the ESXi console, use the following command::

```
# esxcli storage nmp psp list
```

For more information on VMware PSA and NMP, see VMware Knowledgebase article 1011375 or the vSphere Storage Guide.

Capacity considerations

Storage capacity must accommodate VM or differential file growth, data growth, changing business requirements, and the addition of new application needs such as data warehousing to the storage environment.

XP7 storage offers valuable virtualization capacity options, such as storage tiering, to tailor capacity requirements. Custom volume creation then allows specific hard disk drive technologies to support the capacity requirements needed for such things as mixed workloads and hosting golden images.

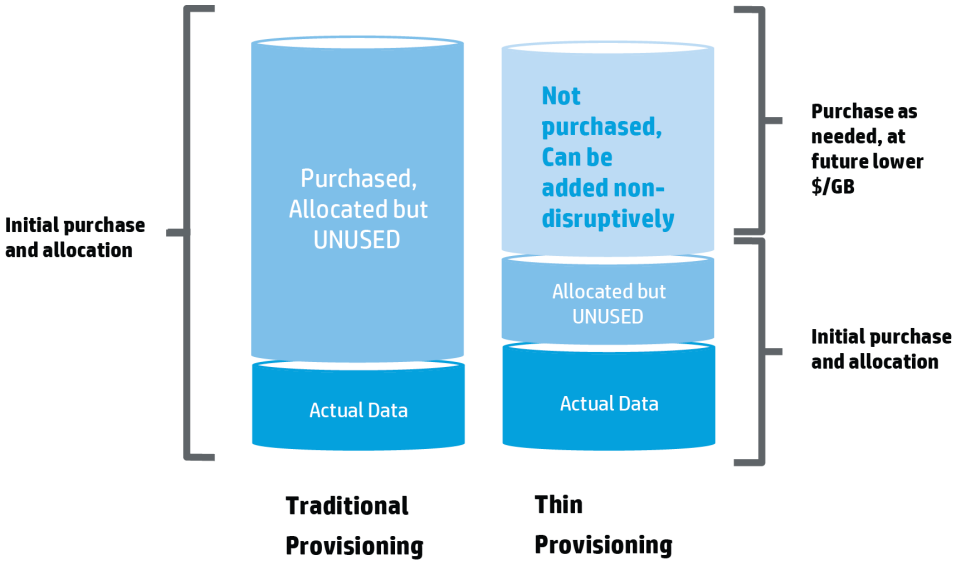
Since XP7 arrays are fully compatible with VMware VAAI, the XP7 offloads processor-intensive tasks from the vSphere host (such as copying hundreds of VM images for VM provisioning) to the storage array, where they are managed more efficiently, conserving host server and network resources. This translates into more efficient storage use.

Thin Provisioning on XP7

System administrators typically provide themselves with more storage than is needed for various applications because they plan ahead for growth. ThP solved the customer's need to better allocate storage based on capacity used, by virtualizing storage to reduce costs and gain efficiency. The introduction of SSD has created additional complications for administrators in sizing for performance with regards to total cost of storage. Using traditional provisioning techniques, the administrator may have to dedicate an excess amount of SSDs to applications in order to satisfy performance and capacity service level agreement .

Thin Provisioning is a technology that presents to hosts large virtual volumes, which are backed up by a pool of physical storage as shown in Figure 9. Smart Tiers expands on Thin Provisioning by automatically migrating data to the most appropriate disk, based on the application's performance needs. A Smart pool is a Thin pool, but with the auto-tiering function enabled.

Figure 9. Comparison between traditional storage provisioning and thin-provisioning on HP XP7 Storage array

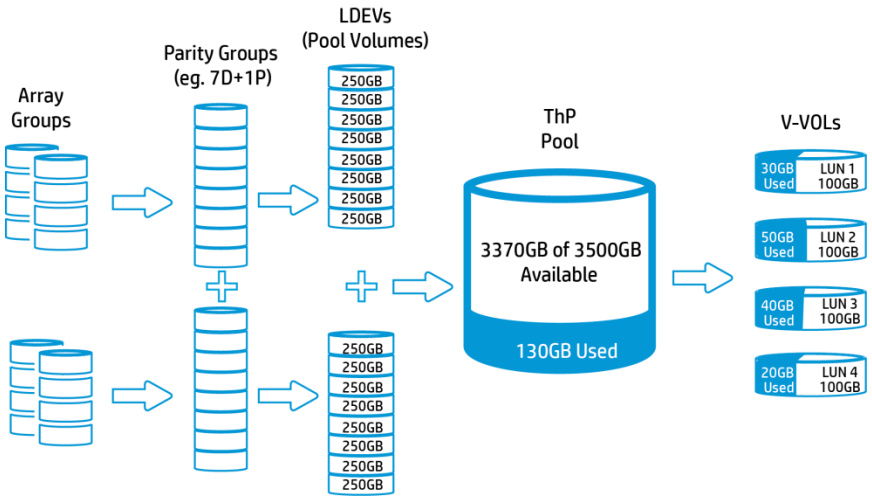


The components of the Thin/Smart solution are:

- Pool Volumes—LDEVs which make up the pool of actual storage.
- Thin Pool—The aggregate of the pool volumes. You can mix different disk types and raid levels, but it is not recommended unless you absolutely need the capacity or plan on enabling Smart Tiers.
- Smart Tiers Pool—A Thin pool with different speeds of pool volumes and the ability to auto migrate data from one tier to another, based on usage.
- Thin Volume—The virtual volume presented to the host, which appears to have much more capacity than is actually the case. Pages of actual storage are allocated from the pool as needed to accommodate writes to the volume. A Smart Tiering-enabled thin volume also has its data distributed across disks of varying speed to optimize its performance.
- Threshold—The pool has a fixed threshold at 80% to warn that its capacity limit is approaching. You can configure an additional threshold to receive earlier or later warnings.
- Subscription Ratio—You can configure how much to over-or under-provision the pool. For example, set the subscription ratio to 100% so that you cannot over-provision the pool.

Pool performance increases as you add more array groups, because the data will effectively wide stripe across more spindles. We recommend using only array groups of the same RAID level and disk speed so your performance is predictable. However, there is no restriction on adding LDEVs from different RAID levels or disk type because you might need to expand capacity in case of emergency and no other disks are available. You can shrink a pool at any time assuming there is enough space left in the remaining disks. However, with current firmware you cannot shrink the head LDEV of the pool. It is possible to configure a pool to be equal in size to the total virtual capacity to avoid over-provisioning. Adjusting the subscription ratio allows you to take advantage of wide striping data, yet not worry about over-allocation.

Figure 10. The mechanics of how storage thin-provisioning works on the HP XP7 Storage array



VMware vSphere's VMDK thin provisioning is an effective method to use the VMFS datastore efficiently without trapping capacity to a VM that is not using all of it. The XP7 also enables thin provisioning by presenting volumes (V-VOLs) from a large shared pool of LDEVs. This section assesses the abilities of each thin provisioning solution and the use of both technologies together.

Thin provisioning with the XP7 not only provides capacity usage benefits, but also simplifies the way the XP7 storage is provisioned for an vSphere host. It is important to understand how thin provisioning pools are constructed on the XP7.

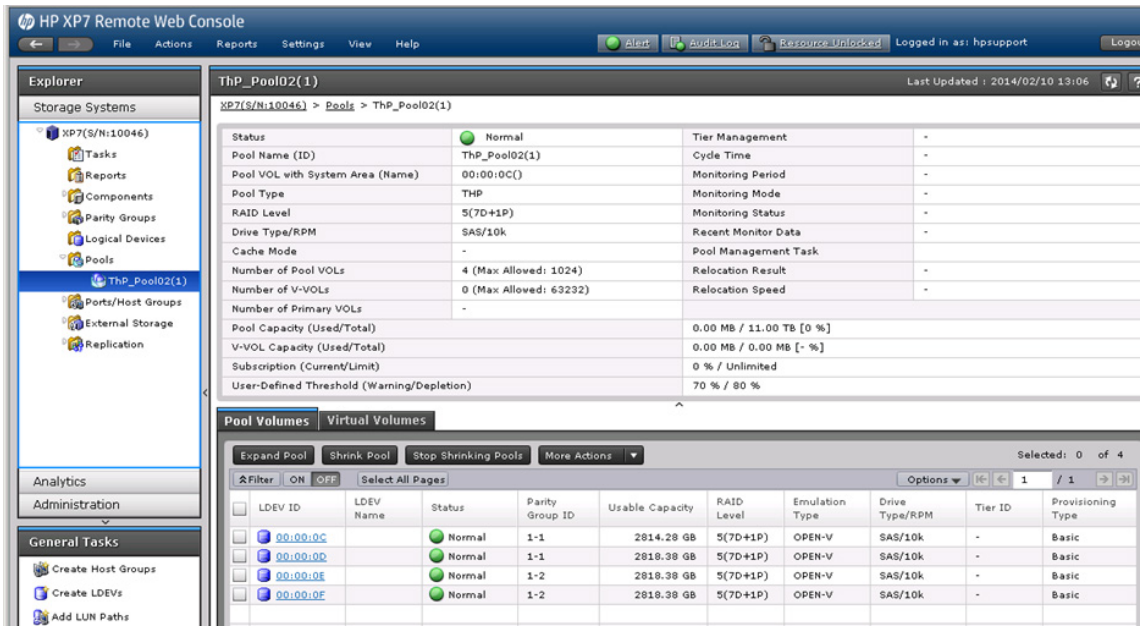
As discussed in the Test environment section, the XP7 begins disk configurations with four disks in an Array Group. Each Array Group is then configured in a RAID set called a Parity Group. After a Parity Group is configured, it is divided into LDEVs with an OPEN-V emulation. LDEVs are then presented to a thin provisioning pool instead of a host as a LUN. LDEVs in a ThP pool are called Pool Volumes. After Pool Volumes have been added to the ThP Pool, capacity can then be presented as ThP Volumes or V-VOLs. The following sections use both of these terms since the figure uses V-VOLs, it would be clearer to the reader if we stuck with one term.

XP7 Thin Provisioning best practices

When creating LDEVs to be used as Pool Volumes, create the same number of equally-sized LDEVs as there are data drives in the Parity Group. For example, if the Parity Group is configured as 7D+1P, divide the Parity Group into seven LDEVs to be used as Pool Volumes.

- Use as many Parity Groups as possible for the ThP pool. Doing so boosts performance because of the number of physical disks and DKAs used.
- Create separate control units (CU) for ThP volumes so they can be easily identified at the vSphere host by their NAA identifier. See Create Host Groups and LUN Mapping section for information on NAA identifier.
- Add the entire Parity Group to the ThP Pool, and do not use any remaining capacity for normal volumes or other pools.
- Set up SNMP agents or email alerts so you receive threshold notifications.
- Create pools based on performance requirements and capacity requirements. Consider the type of hosts that use the storage, such as databases, file servers, and web servers.
- Use the same Parity Group types, spindle sizes, spindle speeds, disk type (SATA or FC), and internal/external storage in a ThP Pool. Mixing these types can cause performance to become unpredictable and inconsistent.
- Be careful when using I/O test tools, defragmenting, or restoring image backups. Because they may cause the entire V-VOL to be fully used. These tools try to touch or restore all blocks in the raw device or on the file-system.
- Create a ThP Pool with RAID 1 Pool Volumes to use ThP for highly available hosts, such as LDEVs from a 2D+2D or 4D+4D Parity Group. To use ThP for lower tier servers or test environments, consider RAID 5 Pool Volumes (LDEVs from a 3D+1P or 7D+1P or an interleaved RAID 5 set like 14D+2P).
- Try not to use V-VOLs for VMFS volumes that host VMware Fault Tolerant VMs. Doing so formats the VMDK drives with *eagerzeroedthick*, and fully allocates the provisioned storage, even though it is not using some of that space.

Figure 11. A ThP pool consisting of four logical devices (LDEV) from two parity groups on an HP XP7 storage array



The figure shows a ThP pool created with two parity groups, as recommended as a best practice.

Scenarios that use both XP7 Thin Provisioning and vSphere Thin Provisioning

After considering both vSphere thin provisioning and XP ThP, determine which solution should be used. Many factors are involved when an enterprise chooses a solution. Each situation requires different considerations and results in different solutions; however, it is not out of the question to consider using both together.

Use vSphere Thin Provisioning with XP ThP for these reasons:

- Use vSphere thin provisioning to use VMFS space efficiently. Use XP ThP for the ThP Pool performance and manageability benefits without over-allocating the XP ThP Pool capacity.
- Use vSphere thin provisioning and XP ThP to gain the advantage of over-allocating storage to gain higher levels of storage efficiency.
- Because there is currently no integration between vSphere thin provisioning and XP ThP, tracking actual disk usage vs. allocated disk usage can become difficult. If both forms of thin provisioning are used, follow these best practices:
  - Keep V-VOLs hosting vSphere thin provisioning in a separate XP ThP Pool.
  - Set up alerts (email or SNMP traps) for both the vSphere Server and the XP7, so administrators are notified well in advance of the disks becoming fully utilized.

Monitoring growth rates

When using vSphere thin provisioning or XP7 ThP, it is important to understand the host volume rate of growth. Understanding growth rates can help in addressing growth needs with plenty of time before all of the pool space is consumed. If the production environment growth rate is unknown, consider using the threshold alerts as markers for growth. For example, set one of the utilization thresholds to 20% and after that trigger is reached, set it again to 25%, repeating this exercise until there are enough data points to map out growth over time. Be sure to keep a watchful eye on ThP Pool usage on the RWC. Use SNMP traps, or set up email alerts to know when capacity limits are being reached. For more information on XP7 Thin Provisioning and best practices, go to the Resource Library on the XP7 product page.

Performance and ThP

ThP is a capacity feature and a good way to enable consistent and enhanced performance. More and more HP customers are using XP ThP as a way to reduce storage administration complexity and attain high performance. Reducing latency and improving transactional performance is best achieved by using an appropriate number of physical disks, and the more there are the better. Interleaved parity groups are not the only way the XP7 can put a large amount of disks behind a single LUN. ThP Pools can contain vast amounts of Array Groups behind a single LUN by putting the LDEVs from these Array Groups into a large, virtualized pool called a ThP Pool.

Tip
Using vSphere thin provisioning with high performance, critical databases like Microsoft® SQL Server, Microsoft Exchange, and Oracle may not be suitable as these solutions need to be more predictable and may incur some performance contention as a result of being on a shared VMFS volume. HP recommends putting application database volumes in a VMFS that is not shared with other VMs that have thin provisioning enabled as this may lead to an unexpected over allocation of the VMFS, causing unpredictable results for the database server.

For more information about provisioning with the HP XP7 Storage system, refer to the *HP XP7 Provisioning for Open Systems User Guide*, which can be found at hp.com/go/XP7.

Note
VMware Site Recovery Manager (SRM) does work with XP ThP. Refer to section "Disaster Recovery with vSphere Site Recovery Manager and HP XP7 Storage."

Array Thin Provisioning and VMFS Datastores

A VMFS datastore that is deployed on a thin-provisioned LUN can detect only the logical size of the LUN. For example, if the array reports 4 TB of storage while in reality the array provides only 2 TB, the VMFS datastore considers 4 TB to be the LUN's size. As the datastore grows, it cannot determine whether the actual amount of physical space is still sufficient for its needs.

However, when you use the vSphere Storage APIs for Array Integration, the host can integrate with physical storage and become aware of underlying thin-provisioned LUNs and their space usage.

Using thin provision integration, your vSphere host performs these tasks:

- Monitors the use of space on thin-provisioned LUNs to avoid running out of physical space. As your datastore grows, or if you use Storage vMotion to migrate VMs to a thin-provisioned LUN, the host communicates with the LUN and warns you about breaches in physical space and out-of-space conditions.
- Inform the array about the datastore space that is freed when files are deleted or removed from the datastore by Storage vMotion. The array can then reclaim the freed blocks of space.
- In the event of an out-of-space condition on the datastore, Thin-Provisioning Stun (VAAI) pauses the affected VMs. This allows the administrator to respond to the Thin Provisioning Space Threshold condition.

Requirements for HP XP7 ThP with VMFS datastores

To use the thin provision reporting feature, your host and storage array must meet the following requirements:

- vSphere version 5.0 or later.
- Storage array has appropriate firmware that supports T10-based vSphere Storage APIs for Array Integration (Thin Provisioning).

Identify Thin-Provisioned storage devices

The `esxcli` command is used to verify whether a particular storage device is thin-provisioned. The command can be run from an SSH session to the vSphere host. In this example, **`--server=`*`server_name`*** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file are supported. For more information on `esxcli` commands and their related options, see Getting Started with vSphere Command-Line Interfaces or the reference vSphere Command-Line Interface Concepts and Examples.

Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. For troubleshooting, run `esxcli` commands in the ESXi Shell or enable SSH on the vSphere host server, under Troubleshooting Options. See VMware Knowledgebase article 1017910 for information on enabling SSH or Tech Support mode on a vSphere 5.x host.

Tip
To validate whether a storage device such as an HP XP7 Storage system LUN is thin-provisioned , use the following command (D=device_ID):

```
# esxcli --server=server_name storage core device list -d naa.XXXXXXXXXXXX4c
```

For more information on VMware vStorage APIs for Array Integration (VAAI), refer to the VMware Knowledgebase article FAQs for VAAI and the VMware Community VAAI document.

For more information on using the `esxcli` commands for identifying disks or LUNs from a vSphere host, refer to VMware Knowledgebase article KB1014953.

vSphere 5 Thin Provisioning

When creating a new VM or adding a hard disk to a VM, you can use one of these three provisioning techniques as shown in table 3. In the VM Creation screen, the default disk type is provisioned *lazyzeroedthick*. By selecting the Thin Provisioning option, the disk is created by using only the amount of capacity equal to the data that is actually written, while presenting what appears to be the fully allocated disk size to the VM. Thin provisioning leaves the unused portion of the disk size as available so that other VMDKs on the VMFS datastore may use it. Selecting vSphere clustering features such as Fault Tolerance creates a hard disk with a provisioning type of *eagerzeroedthick* and disables the ability to create a thin VMDK.

Table 2. vSphere Virtual Disk Provisioning types

Thick Provision Lazy-Zeroed	Creates a virtual disk in a default thick format; space required for the virtual disk is allocated upon creation; storage not zeroed until first write.
Thick Provision Eager-Zeroed	Type of thick virtual disk format that supports vSphere clustering features such as Fault Tolerance. Space required for virtual disk is allocated and all blocks are zeroed upon creation. This type takes more time for virtual disk creation.
Thin Provision	This format is used to save storage space. The thin disk starts small at first, using only as much datastore space as needed for initial operations of the VM.

For more information about virtual disk provisioning, refer to the VMware vSphere 5.5 documentation.

Fault Tolerance in vSphere requires that the entire disk is fully allocated and zeroed. On creation, it writes zeroes to all blocks of the disk.

When using thin provisioning for VMDKs and the VMFS has been safely over-allocated, it is important to monitor VMFS usage to be sure that the VMFS does not run out of space.

Thin Provisioned Block Space Reclamation (UNMAP)

With vSphere 5.5, a brand new VAAI primitive has been introduced, which is called Block Space Reclamation, as part of the overall thin provisioning primitive. Block Space Reclamation is also sometimes referred to as UNMAP. This feature allows for the reclaiming of blocks of thin-provisioned LUNs by communicating to the storage array that specific blocks containing remnant data from VMs which have been moved or deleted are no longer in use. The command used for this operation is the SCSI 'unmap' command. This feature is of great benefit in reclaiming the remnant block space resulting from VM virtual disks which have been migrated or deleted from their original source datastore volumes. The blocks associated with the source VMFS volume are unmapped, thus reclaimed and become usable free space.

The UNMAP primitive is used by the ESXi host to update the Storage Array about the storage blocks that have to be reclaimed after deleting a VM or migrating it to another datastore using Storage vMotion. In vSphere 5.5 # esxcli storage vmfs unmap command is used.

First we need to confirm if the UNMAP VAAI primitive is enabled for block space reclamation on the HP XP7 Storage. Determine this by running the following command:

```
esxcli storage core  device vaai status get -d  <naa>
```

The storage device displays **Delete Status** as supported, meaning that it is capable of sending SCSI UNMAP commands to the array when a space reclaim operation is requested.

Tip

To reclaim unused storage blocks on a VMFS datastore for a thin-provisioned device, run the command:

`esxcli storage vmfs unmap --volume-label=`*volume_label*/--volume-uuid=*volume*_uuid --reclaim-unit=*number*

To verify that the UNMAP primitives are being issued to the vSphere host, run ESXTOP. Press 'u' to get into the disk device view, then press 'f' to change columns, and press 'o' and 'p' to select display "**VAAISTATS**" and "**VAAILATSTATS/cmd**" fields. Monitor the values under "DELETE", "DELETE_F" & "MBDEL/s" columns during a space reclaim operation.

For more information on using the command for Thin Provisioned Block Space Reclamation, refer to VMware Knowledgebase article KB2057513. The ESXTOP utility is covered in more detail in the Using ESXTOP in vSphere 5.5 section of this white paper.

Presenting storage volume to vSphere hosts

Use the HP XP7 Remote Web Console to define paths between hosts and volumes (LUNs) in the disk array. This process is also called 'LUN mapping.' In the Remote Web Console, LUN mapping includes:

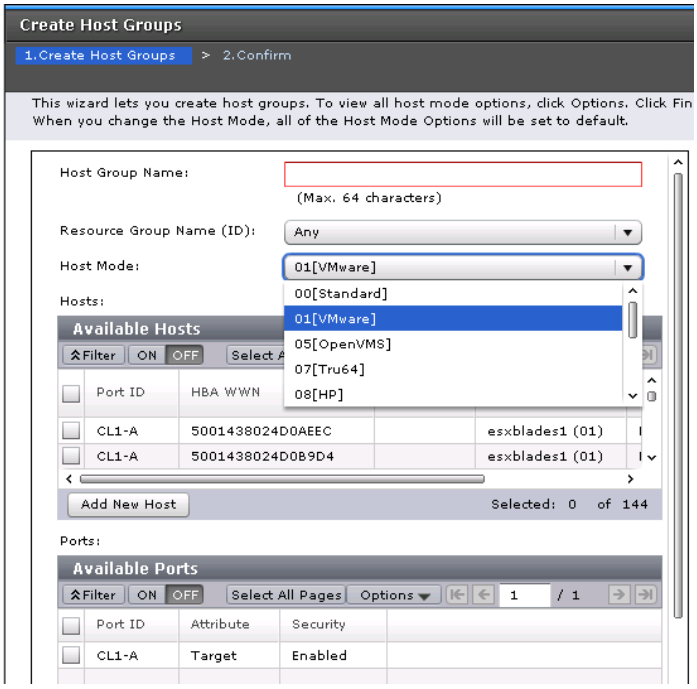
- Configuring ports
- Enabling LUN security on the ports
- Creating host groups
- Assigning FC adapter WWNs to host groups
- Mapping volumes (LDEVs) to host groups by assigning LUNs

Refer to the HP XP7 Storage Configuration Guide for further information. Visit the HP Support Center website.

Create host groups and LUN mapping

Begin the process by selecting the XP7 port to which the vSphere host servers are attached via FC fabric. Create a host group with a unique name. The mode of the new host group must be set to 01{VMware}. See figure 12.

Figure 12. Creating a host group in XP7 Web Explorer



The WWNs of each vSphere host server HBA connected to that port must be selected and added. You can consider this as array-side zoning. A good practice is to open the zone administration window to the FC switch connected to that specific port on the XP7 array. This provides easy reference while creating the new host group and easy WWN verification. See figure 12.

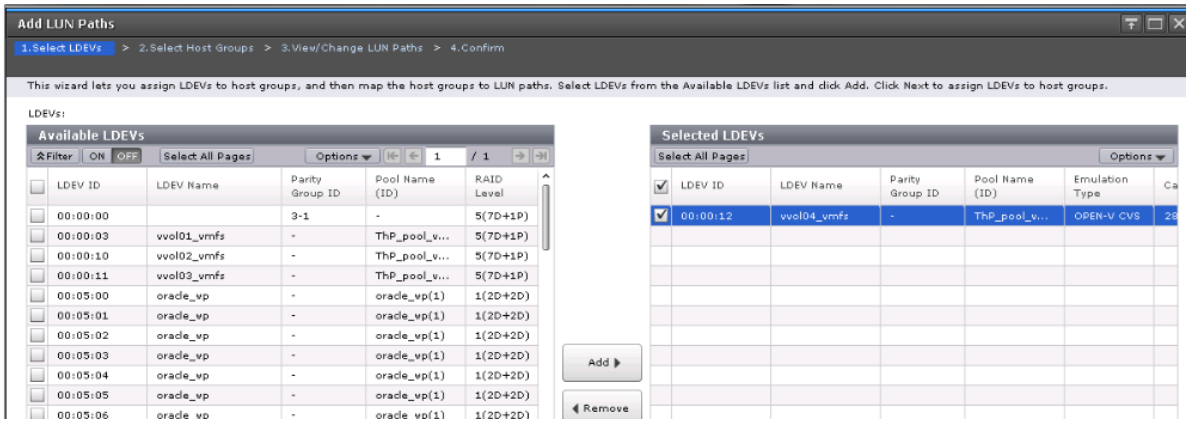
For more information regarding XP7 host groups, see the Host Groups and LUN Mapping section. Also refer to the section "adding resources to a resource group" in the HP XP7 Provisioning for Open Systems Guide.

Caution

Changing the host group modes for XP7 ports in active use by host servers already configured is disruptive and requires the server to be rebooted.

After LDEVs have been created from the ThP pool and host groups have been created with WWNs of the vSphere host server HBAs, the provisioned LDEVs can be mapped to the host group.

Figure 13. Creating a LUN mapping to a host group in XP7 Web Explorer; note the LDEV name

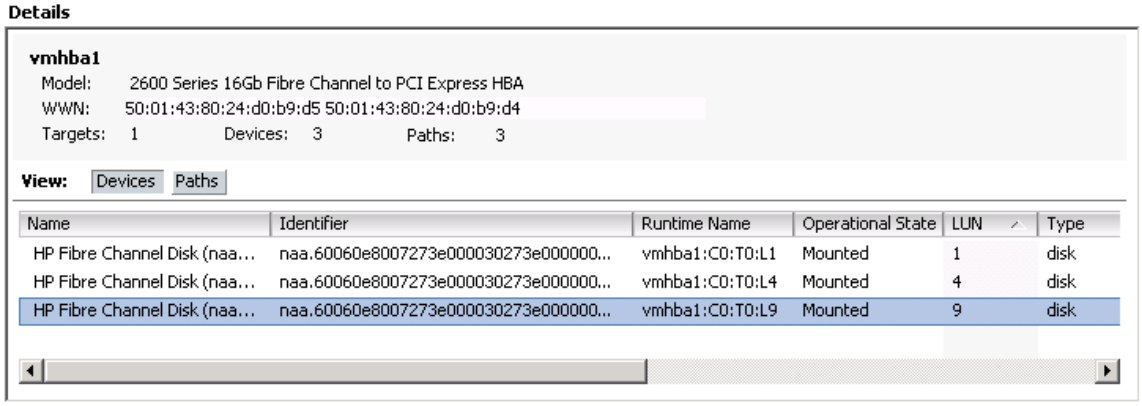


Within the LUN mapping, a specific LUN ID can be selected, which can help in identification of the new LUN from the vSphere client. Selecting a specific LUN ID is an option during the Add LUN Paths process in HP XP7 Remote Web Console. It is a recommended best practice to use the same LUN ID for the LUN mapping on all of the vSphere hosts in the same cluster.

After the LUN mapping has been defined, the new LUN is detected by the vSphere host. VMware vSphere obtains the 128-bit Network Address Authority (NAA) from the extended vital product data in the LUN Inquiry; the NAA identifier provides the HP XP7 Storage array serial number, control-unit (CU) number and LDEV number. The device name seen under Storage Adapters in the vSphere client is automatically generated when the vSphere host detects the new storage (see figure 14); this device name can be changed by the vSphere administrator to a more descriptive name.

Once the LUN mapping has been completed and the vSphere host storage adapters scanned and the new LUN detected, a new VMFS datastore can now be created.

Figure 14. Storage devices (HP XP7 Storage) in vSphere and NAA identifier example as seen in vSphere client



Tip
For ease of identification during the LUN mapping process, it is recommended to label the LDEVs/v-vols in your ThP pool with user-friendly names.

Example: *vvol01_vmfs*

For more information about the LUN mapping steps, refer to the *HP XP7 Provisioning for Open Systems Guide* and the *HP XP7 Remote Web Console User Guide*. These and other HP documents can be found on the following HP web site: hp.com/go/xp7.

VMware multi-pathing considerations

To maintain a constant connection between a vSphere host and its storage, vSphere supports multi-pathing. To take advantage of this support, virtual volumes should be mapped to multiple paths to the host server. To do this, see the previous section on creating Host Groups. The Host Groups include the World Wide Names (WWNs) of multiple HBA ports on the vSphere host server. For an vSphere cluster, the LUNs must be exported to all of the host definitions for the cluster nodes, or a host set may be created containing all of the servers and the VLUNs can be exported to the host set.

VMware vSphere 5.5 includes active/active multipath support to maintain a constant connection between the ESXi host and the HP XP7 Storage array. Three path policies are available: “Fixed”, “Most Recently Used,” and “Round Robin”. For HP XP7 Storage, Round Robin is the recommended policy for best performance and load balancing; however, it may not be enabled by default. The path policies can be viewed and modified from the vSphere Web Client on a per datastore basis as follows:

5. In the vSphere Client or vSphere Web Client, select the datastore.
6. Select the Manage tab, then the Settings tab, and then click on Connectivity and Multipathing.
7. Select one of the vSphere hosts and then click the Edit Multipathing button.
8. In the pop-up window, select Round Robin from the Path Selection drop-down menu.
9. Click the OK button to save the new setting.
10. Repeat steps 3 through 5 for each vSphere host.

Using round-robin and setting the default SATP for the XP7 disk array
The default algorithm for selecting the next path when using Round-Robin is based on the IOPs (input/output operations per second) type as follows:

- IOPs = 1000 (default type)
- Bytes = 10485760 (10 MB)
- Use ANO = 0 (active, non-optimized is an ALUA setting which is not used with the XP7)
- I/Os Pending = 0
- Bytes Pending = 0

In summary, for every 1,000 operations, select the next path. This default algorithm can be changed. Try using the following command to understand more about changing the default algorithm parameters, which can be done per device.

```
# esxcli nmp roundrobin setconfig –help
```

After making the changes, the settings can be confirmed with the following command, where naa.X is the full naa name of the given LUN (for example, naa.60060e8005278c000000278c00000090).

```
# esxcli nmp roundrobin getconfig –device naa.X
```

There are many advantages to using Round-Robin with XP7 that will help save time in configuration and deployment, and improve performance and availability. While there are many benefits, it is important to understand the load that will be applied on the disk array and to understand the effect that this multi-pathing policy has on both the host and the disk array, especially in environments with high I/O demands.

In most cases, the experience should be positive; however monitor read-ahead cache hits to make sure there is an even distribution of I/O across all active paths. Read-ahead cache hits are notable because when a sequential read stream is detected by the XP7, it can pre-fetch increasing amounts of data, up to 32 MB. The more sequential the workload is, the more the XP7 will pre-fetch into cache. If the HP XP7 Storage is used with VMware vSphere 5.5, and the VM perform heavy read I/O, as in a data-warehouse type of application, pay attention to any indicators of sequential percentages and cache hits. As a sanity check, compare the performance to that of a FIXED path performance. Testing conducted in support of this white paper showed that there was no performance impact on heavy read I/O, and that the sequential access was the same when using either FIXED or Round-Robin.

If Round-Robin is the path policy of choice for a vSphere 5.5 and HP XP7 Storage implementation, the default behavior of a new LUN to the system can be changed. Before changing the SATP PSP assignment for HP XP7 Storage, all new LUNs to the vSphere host will have a path policy of FIXED, which means an administrator must balance the fixed path or set the LUN load balancing policy to Round-Robin. This procedure must be conducted for every new XP7 LUN on the host. Changing this default behavior can save storage administrators time and can be done by using the following command in the ESXiconsole CLI.

```
esxcli nmp satp setdefaultpsp --psp VMW_PSP_RR --satp VMW_SATP_DEFAULT_AA
```

After executing this command, all new LUNs from the XP7 Disk Array will have an applied Round-Robin path policy.

Note
If the vSphere environment is intended to include VMs which will be clustered via the Guest OS, such as Microsoft Cluster Service (MSCS), the multipath policy should be considered. In vSphere 5.5, round-robin (PSP_RR) is now supported for VM disks engaged in MSCS clustering. The following restrictions apply:

- Supported on Windows 2008 and Windows 2012 server guest operating systems only.
- Supported on Cluster Across Boxes (CAB) and N+1 configurations. Cluster in a box uses Virtual Reservations.
- Shared disk (Quorum OR Data) must be provisioned to guest in pass-through RDM mode only.

To avoid unnecessary cluster node failovers due to system disk I/O latency, virtual disks must be created using the EagerZeroedThick format on VMFS volumes only, regardless of the underlying protocol. Configurations using shared storage for Quorum and/or Data must be on FC based RDMs (physical mode for cluster across boxes "CAB", virtual mode for cluster in a box "CIB") in vSphere 5.1 and earlier. Windows Server 2012 failover clustering is not supported with vSphere-provided shared storage (such as RDMs or virtual disks) in vSphere 5.1 and earlier. VMware vSphere 5.5 provides complete support for 2012 failover clustering. Refer to VMware Knowledgebase articles 1037959 and 2052238 for further information regarding support for MSCS at the VM layer.

VMware and storage performance

Highly available and high-performance storage and server solutions are an integral part of the goal of an always available virtualization environment. With the vSphere vStorage API integration of the HP XP7 Storage system, VMware vSphere 5.5 is more efficient in using the XP7 Storage, reducing resource usage on the host server hardware platform.

For more information on VMware vSphere storage performance considerations, refer to the VMware vSphere 5.5 Documentation Center, and the VMware vSphere blog—Troubleshooting Storage Performance in vSphere.

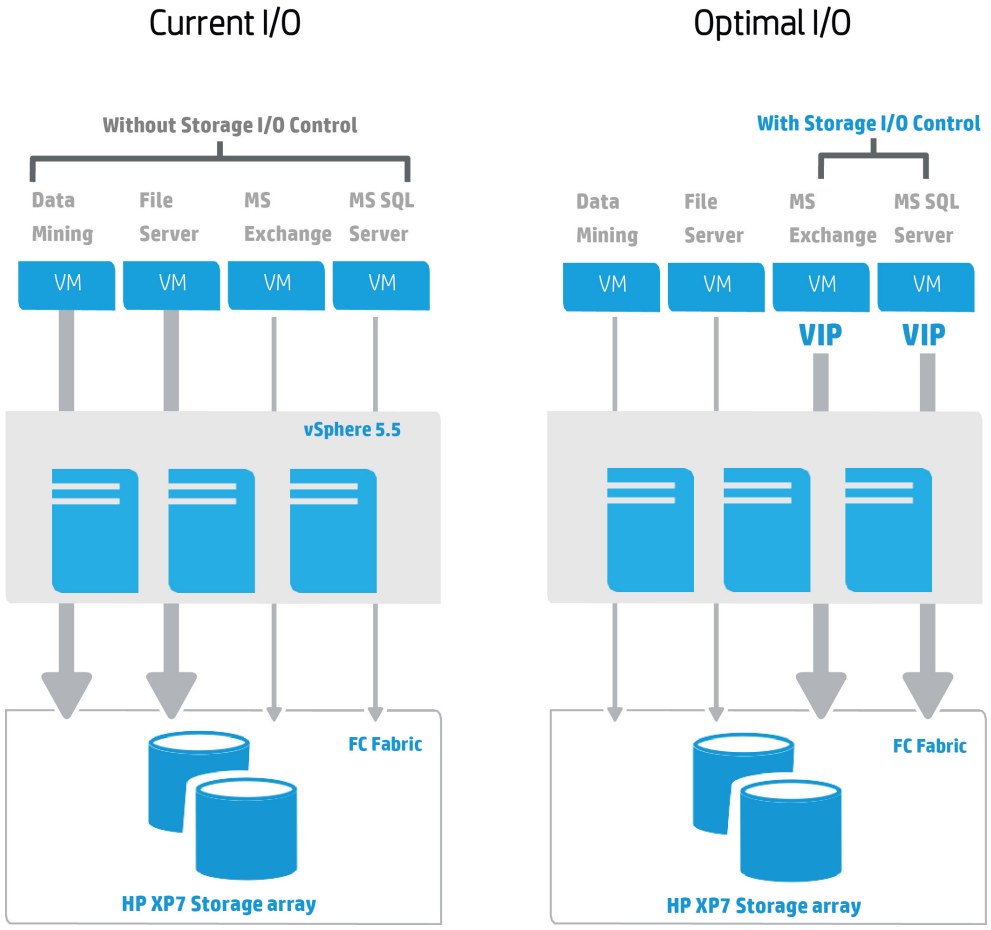
Storage I/O Control (SIOC)

The vSphere Storage I/O Control (SIOC) feature manages shared storage resources across vSphere hosts to provide more predictable performance during periods of congestion. It monitors the latency of I/Os to a datastore for each vSphere host sharing the device. When the average latency for a datastore exceeds a threshold set by SIOC based on storage type, SIOC distributes the storage resources to VMs according to their assigned shares. It accomplishes this by reducing the number of I/O queue slots available to lower priority VMs and increasing the slots for VMs with higher shares. By controlling the vSphere device-queue depths in proportion to the VM shares, SIOC is able to control storage congestion for the datastore and distribute HP XP7 Storage array resources appropriately.

The congestion threshold represents a tradeoff between lower I/O latencies and throughput. When the threshold is set low, I/O throttling is engaged more aggressively, which helps to maintain a lower I/O latency for the datastore, but also reduces the overall throughput for the datastore.

The default latency threshold is 30 ms, but the optimal setting depends on the storage type used for the datastore. A new feature in SIOC for vSphere 5.1 is the I/O injector, which calculates the peak throughput for a storage device, detects the 90 percent throughput value and measures latency at that point to determine the optimal threshold value for the storage device. The latency threshold is set automatically to the value determined by the I/O injector, but you can change the 90 percent throughput value or set the threshold in milliseconds if desired.

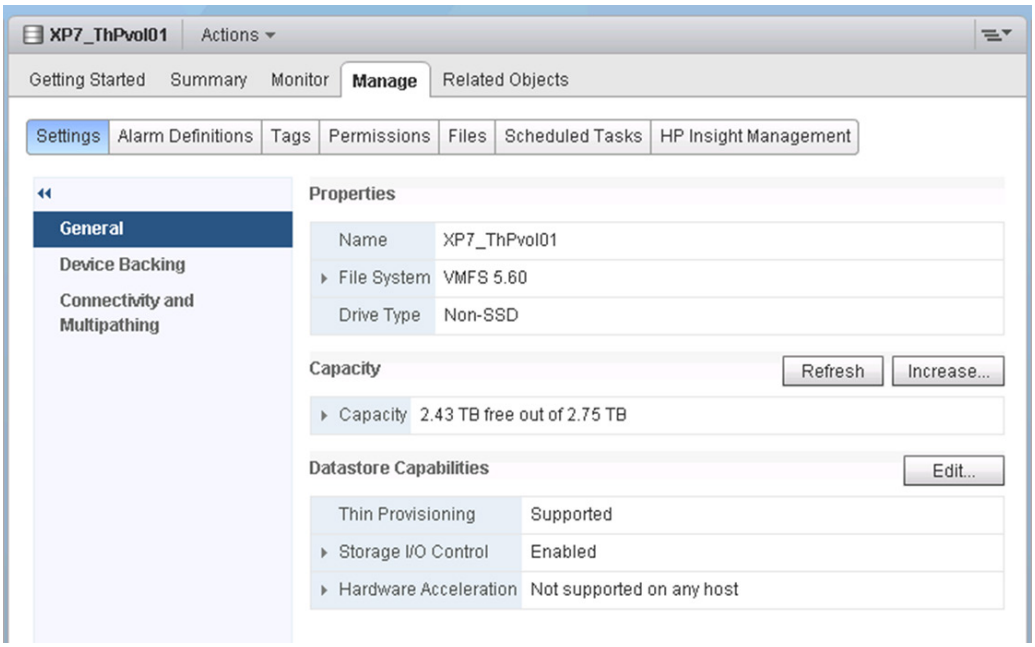
Figure 15. Comparison between VM I/O without SIOC and VM I/O with the benefit of SIOC enabled. The 'VIP' VMs are assigned higher SIOC shares



To enable SIOC for a datastore:

1. In the vSphere Web Client, select the datastore.
2. From the **Manage** tab, select Settings and click the Edit button next to Datastore Capabilities.
3. Check the **Enable Storage I/O Control** check box.
4. If you want to change the peak throughput threshold, open the percentage of "peak throughput drop down list", to select a new value.
5. Alternatively, set the threshold to a specific latency by clicking the **Manual** radio button. Select a value in milliseconds from the drop-down menu.

Figure 16. Datastore detail in vSphere web client, with SIOC setting enabled



The VMware white paper Storage I/O Control Technical Overview and Considerations for Deployment makes the following recommendations for determining the congestion threshold values for various types of storage, including auto-tiered storage. However this source is based on version 4.1 of vSphere and does not reflect enhancements in SIOC introduced since vSphere 5.0 was released.

Table 3. Recommended congestion threshold values for SIOC

Type of storage/disks in array	Recommended threshold
SSD	10-15 ms
Fast Class (FC)	20-30 ms
Nearline	30-50 ms
Auto-tiered storage	Combine ranges of fastest and slowest storage types

The default threshold of 30 ms should be acceptable for most workloads using Fast Class or Nearline drives. If Smart Tiers feature is employed, then the threshold should be set to a value that is within the recommended ranges for the fastest and slowest storage types in use. For example, if Fast Class and SSD drives are specified by the Smart Tiers policy, then the threshold for the datastore should be set in the range of 10 to 30 ms. The exact setting should be determined by the requirement to favor low latency (using a low congestion threshold) vs. throughput (using a high threshold).

The VM's relative priority on the datastore is determined by its number of disk shares. This is configured by editing the settings for the VM. Select **Virtual Hardware**, and select the hard disk of interest, the default number of disk shares is 1000. To increase the priority of a VM relative to other VMs using the same datastore, increase the number of disk shares by clicking on the Shares dropdown menu and change the value to **High**, or select **Custom** and enter a new value in the adjacent field.

VmObservedLatency/Storage I/O Control (SIOC) —Measures the time between receipt by VMkernel of the I/O from the VM and receipt of the response from the datastore and will automatically adjust the latency threshold. When running 16GFC and SSD storage, VMware vSphere 5.1 enables lower latency thresholds and delivers higher overall storage I/O performance. This reduces the possibility of I/O congestion that could trigger the SIOC and result in lower performance

For more information about Storage I/O Control feature in vSphere 5, visit the VMware website. Troubleshooting information on SIOC is available in VMware Knowledgebase article KB1022091.

Adaptive queue depth throttling

Each port on the HP XP7 Storage system has a finite queue depth that depends on the HBA model; each server attached to a port shares that port's queue. If a host sends an I/O request to a port with a full queue, the host receives a "queue full" SCSI response from the HP XP7 Storage system. I/O commands sent to a port in an HP XP7 storage array that has reached its maximum queue depth are not processed beyond the "queue full" SCSI response.

Historically, a vSphere host's default reaction to this response would be to recognize it as a valid command and to continue sending requests to that port. Lack of I/O responses can result in VMs becoming unresponsive and can lead to a crash of the vSphere host. ESXi 3.5 Update 4 and later include an adaptive queue depth throttling algorithm which adjusts the LUN queue depth in the VMkernel I/O stack. This algorithm is activated when the storage array indicates I/O congestion by returning a "queue full" SCSI status. When congestion is detected, the VMkernel throttles the LUN queue depth and attempts to gradually restore the queue depth when congestion conditions subside.

Without adaptive queue depth throttling, administrators are forced to limit the number of VMs per physical server to reduce the risk associated with any particular VM overrunning I/O queues. Administrators are also forced to manually tune the number of VMs when they detect congestion—a reactive, slow, and error-prone process. By automating congestion control, administrators can confidently create a higher number of VMs per physical server without the need for manual congestion control.

The adaptive queue depth algorithm is disabled by default. For vSphere/ESXi 5.5, it is enabled on a per datastore basis by setting the queue-full-sample-size and queue-full-threshold parameters. Setting the queue-full-sample-size parameter to a value greater than zero activates the algorithm. The queue-full-threshold parameter must be set to a value less than or equal to queue-full-sample-size. To set these parameters to optimal values for HP XP7 Storage, run the following command for each HP XP7 Storage device utilized by the ESXi host:

```
#esxcli storage core device set --device device_name --queue-full-threshold 4 --
queue-full-sample-size 32
```

These settings take effect immediately and are persistent across reboots of the ESXi hosts. Note that it is important to make the changes across all ESXi hosts sharing the storage.

For ESXi versions before 5.1, the algorithm was enabled by setting two VMware system-wide configuration parameters, QFullSampleSize and QFullThreshold on the ESXi hosts. The new per-device settings are preferred because the optimal settings differ by storage type. For more information, refer to the VMware Knowledge Base entry: kb.vmware.com/kb/1008113.

Note

Adaptive queue depth throttling is not needed if Storage I/O Control is enabled.

Adaptive queue depth throttling is not compatible with Storage DRS.

If adaptive queue depth throttling is enabled, it is important to enable it for all hosts which are accessing the HP XP7 Storage.

XP7 storage performance monitoring

The two main tools to monitor the performance of the XP7 storage array are the XP Performance Monitor and the vSphere CLI tool ESXTOP. Both are useful for measuring the HP XP7 Storage performance under a virtualization workload.

Even though the HP XP7 Storage delivers optimal performance, getting top performance from the vSphere host and VMs can be a challenge. There are multiple possible adjustments and locations in the performance chain where bottlenecks can exist. There are always bottlenecks, and the goal is to identify the bottleneck location and then determine how to remove that location as a bottleneck. The first step in determining performance is to know the theoretical maximum that can be delivered from the array. Array performance is typically a factor of spindle speed, spindle count, RAID overhead, and cache use. The XP7 is a complex array where all of these factors come into play. For more information about sizing the XP7, contact an HP technical specialist.

Aside from adding more spindles and cache, there are other opportunities that have shown to enhance performance. This section covers storage performance with VMware vSphere 5.5.

Before beginning, understand the performance you need to obtain and plan for how those metrics will be obtained. This information is imperative in achieving the desired results with little wasted time, effort, and money. When possible, determine the performance of the physical environment before moving into a virtual server environment, or use HP Storage sizing tools to identify what the goals of a specific server environment will be.

Working with the defaults

Performance with vSphere 5.5 and the HP XP7 Storage is often good out-of-the-box as many of the defaults for these environments are satisfactory. Establish a baseline before performance tuning so that the improvements can be understood and the effort to make the changes can be justified. Change control is recommended when capturing baseline performance data from your vSphere hosts. Refrain from installing updates and patches on the vSphere hosts until after baseline data, performance data, and configuration changes have been recorded.

Monitoring performance

This solution offers many tools for evaluating performance. Located within the VM, at the fibre-channel switches, at the vSphere host layer, and at the HP XP7 Storage. Because any number of operating systems can be hosted by vSphere, and each of those operating systems supports different performance monitoring tools, this section reviews ESXTOP and HP XP Performance Advisor (XPPA) Software. Esxtop is a native tool to vSphere which can monitor network, CPU, memory, and storage performance. XPPA offers performance and capacity measurements within the XP7 disk array. Both tools can run remotely or locally. The vSphere Client is also a good place for higher level observations and can be a good tool for performance analysis.

XPPA

HP XP Performance Advisor Software is a web-based application for collecting, monitoring, and displaying XP7 array performance. With XP Performance Advisor, you choose the time and duration of performance data collection—so you can pinpoint activities that significantly impact your disk array and tune the array accordingly. You can easily monitor storage performance and display usage statistics for your storage system at any level, from a complete system overview down to individual components. You can also filter hosts, arrays, and array components, so that you only see the hosts and arrays that you want to see. You can also generate, schedule, and view detailed performance reports that allow you to identify any performance bottlenecks in your HP XP7 Storage system.

Some of the common collection statistics used in this project are cache usage, writes pending (writes waiting to be de-staged to disk from cache), I/O per second, MB/s, sequential versus random percentages, and response time metrics.

To assist with troubleshooting application performance issues, XPPA allows monitoring of each server (host) and application based on the host group name.

Using ESXTOP in vSphere 5.5

The ESXi tool ESXTOP provides real-time insight on performance directly from the vSphere host, via the ESXi command line in an SSH session.

When ESXTOP monitors storage performance on a vSphere host server, it displays various statistics:

- DISK QUED (disk queue depth)
- DISK DAVG (disk latency most likely caused by storage array)
- DISK ABRTS/s (Aborts issued by the guest OS on VM because the storage is not responding. Can be caused by storage path failure)
- DISK CONS/s (SCSI reservation conflicts per second. Can be caused by lock on the VMFS)
- ATSF (the number of ATS commands failed)
- AAVG/suc (the average ATS latency per successful ATS command)
- ZERO (the number of ZERO_BLOCKS commands successfully completed)
- MBZERO/s (Megabytes zeroed per second)
- ZAVG/suc (the average latency per successful ZERO command)

For more information on ESXTOP, refer to the ESXTOP documentation on VMware's website: communities.vmware.com/docs/DOC-11812

Tip

After ESXTOP has been configured to display or capture the fields or metrics of interest, it can be configured to run in batch mode and output the results to a file. Run the following command from an SSH session to the vSphere hosts:

Esxtop –b –d 2 –n 100 > esxtopcapture.csv

The –b is batch mode and the –n switch toggles 100 iterations. For more information on using ESXTOP for storage performance analysis, refer to the section "Storage Performance Analysis and Monitoring" on the VMware Community website.

Disaster Recovery with vSphere Site Recovery Manager and HP XP7 Storage

Along with the escalating costs of maintaining business-critical applications, for many small to midsize enterprises, the cost of replicating their application servers to recover from disasters can be prohibitive. Implementing a cost-effective data recovery (DR) solution using a VMware-certified SAN infrastructure that supports the failover between vSphere environments and VM instances of each business-critical application helps to control costs and provide business continuity.

The VMs used for DR failover can all reside on a small number of DR servers running VMware vSphere. Although the VM instances may not offer the bandwidth to accommodate peak workloads, the solution avoids downtime. If necessary, you can use VMware Storage vMotion to migrate these VMs to more powerful servers as required. The VMware SRM does not replace the functionality of Remote Snap; rather it provides the mechanism for replicating VM data between the protected site and the recovery site.

The HP XP7 Storage Replication Adapter (SRA) provides a platform for a seamless integration of HP XP7 Storage with SRM and aids in the automated DR of VMs. It enables SRM to coordinate the recovery process with the underlying replication mechanisms, so that the VMs at the protected site are shut down properly (when the protected site VMs are available) and the replicated VMs can be activated. SRA helps in executing the DR plans.

A recovery plan specifies the order in which the protected VMs are to be recovered. The recovery plan also specifies network parameters, such as IP addresses, and consists of user-specified scripts that perform custom recovery actions.

After a recovery, the VMs that are up and running are no longer protected. To reactivate protection, SRM supports a reprotect operation for VMs protected on an HP XP7 storage system. The reprotect operation reverses the roles of the two sites after the original protected site is backed up. The site that was formerly the recovery site becomes the protected site, and the site that was formerly the protected site becomes the recovery site.

VMware SRM also enables recovery plan testing using native functionality, without disruption to production environments, by using a temporary copy of the replicated data. Conduct tests after a reprotect operation to confirm that the new protected recovery site configuration is valid.

HP Insight Control Storage Module for VMware vCenter Server

HP Insight Control Storage Module for VMware vCenter Server is a storage management integration plug-in. HP Insight Control for vCenter provides the ability to create datastores and VMs directly from vCenter Server. HP ICSM does not currently support storage provisioning operations (create storage array devices, map them to array port) on the HP XP7 storage array.

HP Insight Control Storage Module for VMware vCenter performs these tasks:

- Obtain physical and logical end-to-end views of storage devices.
- View reports on device capacity and usage.
- Monitor storage information.

HP Insight Control for vCenter Server saves time and resources, allowing administrators to manage both VMware vSphere and HP Storage from a common interface.

HP Insight Control for vCenter integrates with the vCenter management server (also known as vCenter Server) and the vSphere Client software from VMware. VMware vCenter is the single point of management for VMware virtual environments. These virtual environments can consist of many standalone or clustered vSphere configurations. The virtual environment administrator accesses the vCenter management capabilities using the vSphere Client software.

HP Insight Control for vCenter software can be installed on a management server or a VM and configured to connect and register with a vCenter Server. After registering Insight Control for vCenter with a vCenter server, all vSphere clients connected to the vCenter server can use the HP Insight Software tab to access the Insight Control for vCenter software.

The HP Insight Control Storage Module for vCenter is available online from this HP website and can be installed and used at no charge. You will need to accept the HP software license terms as part of the download and installation sequence. An HP Passport sign-in will be required.

Conclusion and benefits

When supported with the correct underlying storage platform, server virtualization delivers increased infrastructure consolidation, administrative efficiency, business continuity, and cost savings. As a result, server virtualization is not only transforming the data center, but also the businesses that those data centers fuel. However, these transformative results depend on enterprise class storage to deliver the performance, availability, and flexibility to keep up with the dynamic and consolidated nature of virtualized server environments. As an enterprise-class storage solution, the HP XP7 addresses the storage-related challenges of virtualization and exceeds expectations.

Virtualization challenges	**Solution–HP XP7 Storage and VMware**
Depleted resources, which result in diminished performance and availability	HP XP7 Storage Thin Provisioning and ThP pools
Lack of application awareness; OS virtualization doesn't virtualize the application	VMware vSphere 5.5, VMware ThinApp
Additional unexpected costs which cause the virtual solution to cost more than the physical problem	Enterprise-class storage solution, the HP XP7 addresses the storage-related challenges of virtualization and exceeds expectations (no forklift upgrades)
Under-utilized virtualization features due to the network limiting implementation	HP 5900 data center switches, HP Virtual Connect Flex-10 switches for HP BladeSystem, HP LOM (LAN on Motherboard)
Increased storage requirements because VM growth exceeded planning	HP XP7 Storage ThP pools, VMware vCenter and HP Insight Control Storage Module for VMware vCenter
Congested storage network	HP StoreFabric and FlexFabric
Management complexity due to management tools lacking interoperability	HP XP7 Remote Web Console, VMware vCenter, integration with storage via HP Insight Control Storage Module for vCenter

Designed for organizations that cannot afford any downtime, the HP XP7 Storage array combines a high-performance, online scalable, fully redundant hardware platform with unique data replication capabilities integrated with clustering solutions for complete business continuity.

As data center environments are virtualized and consolidated on a single storage platform, the I/O pattern seen by the storage device tends to appear more random in nature than might otherwise be the case, which is yet another reason that HP XP7, with its wide-striping, is a great choice for virtualization and consolidation.

Today's dynamic business world ups the ante for virtualized solutions, and the HP XP7. It provides high-availability, storage efficiency, effortless management, and unsurpassed performance in unpredictable heterogeneous environments. Teaming HP XP7 with VMware vSphere in virtualized environments provides the competitive edge that organizations hunger for today.

Resources

HP XP7 Storage Array
hp.com/go/xp7

HP Support for VMware
HP Unified Sizer for Server Virtualization

HP Insight Control Storage Module for VMware vCenter Technical Documentation

HP Insight Control Storage Module for VMware vCenter Download

HP-supplied vSphere 5.5 image on the HP ProLiant website

HP Virtual Connect FlexFabric Cookbook

HP SAN Design Reference Guide

HP SPOCK website (Single Point of Connectivity Knowledge for HP Storage)

HP StoreFabric Storage Networking

VMware knowledgebase articles
blogs.vmware.com/vsphere/2012/05/troubleshooting-storage-performance-in-vsphere-part-1-the-basics-.html

blogs.vmware.com/vsphere/2012/06/troubleshooting-storage-performance-in-vsphere-part-2.html

blogs.vmware.com/vsphere/2012/06/troubleshooting-storage-performance-in-vsphere-part-3-ssd-performance.html

blogs.vmware.com/vsphere/2012/07/troubleshooting-storage-performance-in-vsphere-part-5-storage-queues.html

vSphere distributed switch best practices paper

Additional VMware vSphere Storage configuration guides
Refer to the VMware Storage Guide for guidance on using vSphere with HP MSA storage, iSCSI, or FC storage, and related best practices not covered in this white paper. For information specific to storage protocols and VMware vSphere, the following resources are available:

- VMware iSCSI SAN Configuration Guide
- VMware Fibre Channel SAN Configuration Guide
- vSphere 5.5 Storage Guide

Learn more at
hp.com/go/storage

Sign up for updates
hp.com/go/getupdated

Share with colleagues

Rate this document

4AA5-2046ENW, April 2014