



HP StorageWorks MPX200

Simplified Cost-Effective Virtualization Deployment

Table of contents

Executive Summary	2
HP StorageWorks MPX200 Architecture	2
VMware architecture	3
MPX200 iSCSI solutions for traditional VMware virtualized environment	3
HP StorageWorks MPX200 Solution—improved implementation methods for VMware	5
Implementation Option 1	5
Architecture	5
Benefits	5
Implementation Option 2	7
Architecture	7
Benefits	7
HP StorageWorks MPX200 Benefits	8
Conclusion	9
For more information	9

Executive Summary

Server virtualization helps to improve the Return on Investment (ROI) by increasing server utilization. However, increased server utilization means a greater demand for I/O bandwidth and more connectivity. To meet the virtualization requirements, Storage Area Networks (SAN) must be capable of providing a higher level of scalable performance than is required by traditional server environments.

The rapid growth of virtual servers enables companies to extract higher returns on their computing resources. Because this growth has accelerated in the past few years, many organizations are running into limitations in performance, flexibility, and scalability in deployment of storage in virtual environments. These limitations have resulted in businesses decelerating their planned deployments, thereby not realizing the complete benefits of virtualization.

In a virtualized environment, the hypervisor manages virtual machines. In VMware, the maximum number of storage logical unit numbers (LUNs) that can be exposed to a physical server, before being subdivided into applications, is 256. This limit applies collectively to all hypervisors within the same storage zone. All hypervisors need access to every LUN to manage the migration of virtual machines (such as, VMotion). When the number of virtual machines deployed increase, the storage requirements grow exponentially, and this limitation increases rapidly.

This results in lack of scalability, flexibility, and causes the following problems for storage administrators:

- Every application has the potential to be exposed to every data LUN.
- Security policies are hard to manage and maintain.
- Growth of application data can cause frequent restructuring of the environment.
- If more than 256 LUNs are needed, the hypervisor subdivides and manages the LUNs for the applications, thereby increasing storage complexity and management.

This white paper describes a cost-effective and easy-to-manage solution that allows the Storage Administrator to leverage the current HP StorageWorks Enterprise Virtual Array (EVA) storage investments for SCSI over IP (iSCSI) connectivity and virtual machine deployments. This white paper describes the VMware deployment options for a scalable virtualized data center.

HP StorageWorks MPX200 Architecture

The HP StorageWorks MPX200 Multifunction Router (MPX200) provides powerful connectivity technology, providing modular multi-protocol SAN designs with increased scalability, stability, ROI, and simpler to manage, secure storage solution for virtualized server environments.

The MPX200 with EVA iSCSI solution addresses the virtualization environment issues in multiple ways:

- More iSCSI GbE ports—four ports in a single blade, and eight ports in a Chassis.
- 10 GbE iSCSI connectivity enables customers to run more virtual machines/applications on a single physical server. 10 GbE provides customers a more compelling reason to adopt iSCSI.
- Scalability up to 600 iSCSI initiators on the iSCSI side with architectural support up to 1,024 iSCSI initiators.
- Up to four EVAs can be connected to the same router to provide iSCSI connectivity. It reduces the cost of connecting multiple arrays to servers with iSCSI connectivity.
- 1,024 maximum supported LUNs enables functional deployment of a large number of iSCSI initiators and overcomes the traditional 256 LUN limitation.

The rich features available in the MPX200 iSCSI solution enable several options to build virtualization data centers and improve or preserve the ROI.

Server Virtualization and SAN based Storage

VMware architecture

The VMware ESX server hypervisor abstracts all the details about the underlying hardware, including SAN storage, and provides a consistent view of the hardware details to the virtual machines (VM). Thus, the hypervisor centrally manages all the hardware details and provides services to VMs. In addition, within the virtualized environment, the VM configuration, Guest OS, applications, and application-specific data are encapsulated as files and managed by the VMware's cluster file system (the virtual machine file system (VMFS)). Traditionally, the VMFS manages VMs, applications, and data. The following section describes how MPX200 can be configured in a traditional server virtualization configuration to maximize the benefits of virtualization.

MPX200 iSCSI solutions for traditional VMware virtualized environment

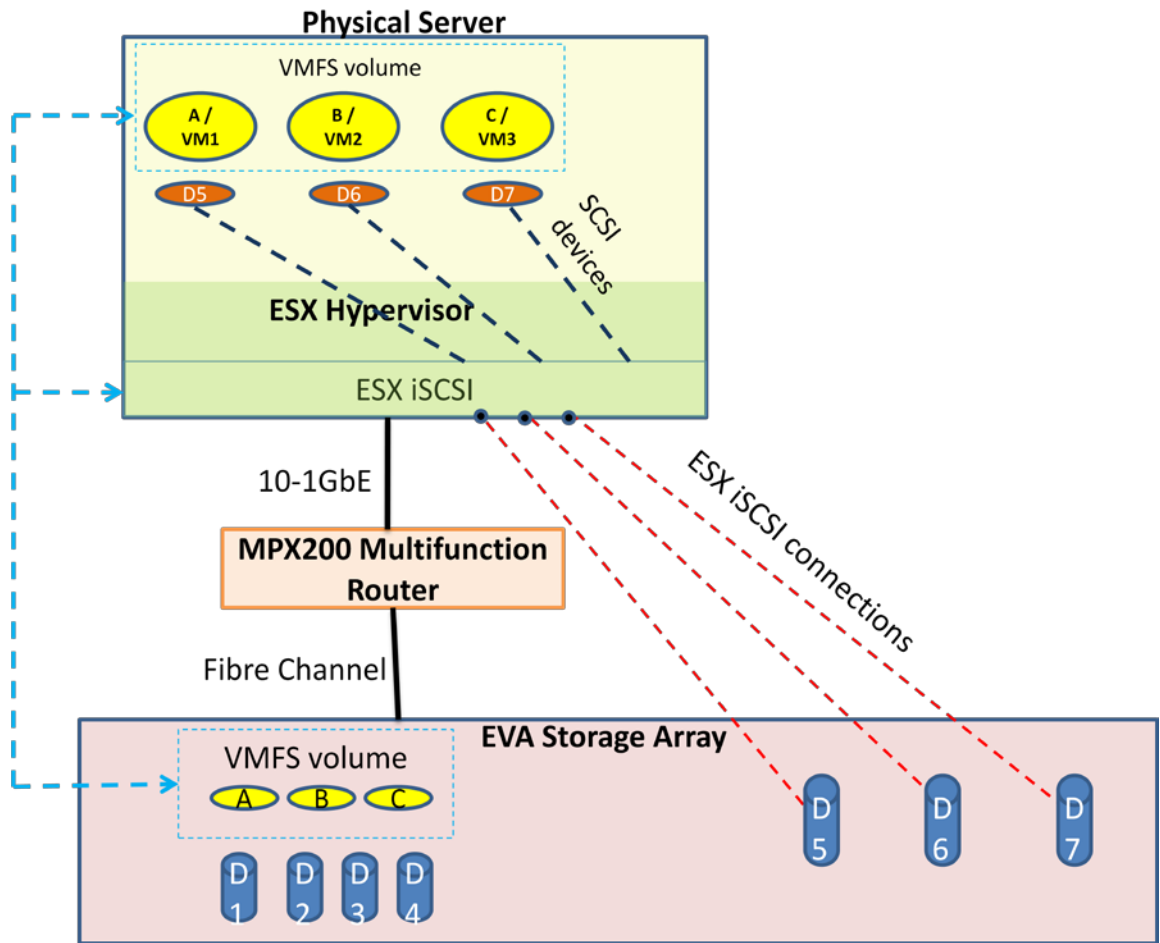
The VMware hypervisor manages storage LUNs as:

- Virtual Machine File System (VMFS)
- Raw Device Mapping (RDM)

Virtual Machines can access storage LUNs by configuring them as VMFS or RDM LUNs. If VMFS is used, the ESX server or hypervisor manages the entire VM and the associated applications/data. If RDM is used, the VM is allowed to manage its own applications and data using its own file system. However, the low-level LUN management remains with the ESX hypervisor only.

[Figure 1](#) shows the traditional way of configuring iSCSI LUNs in a hypervisor and presenting them to virtual machines.

Figure 1: Traditional way of configuring iSCSI LUNS in a hypervisor



Applications running on VMs and data LUNs for these applications may be configured into one of the following options:

- The iSCSI SAN LUNs (EVA Vdisks D1, D2, D3, D4) can be pooled into the VMFS file system volume and presented to VMs (VM1, VM2, VM3) by carving out this volume, using VMFS, as needed individually by the applications running on these VMs. The native file system of the guest operating system can be installed on the VMFS carved-out volume, and the application and data can be stored here.
- EVA Vdisks (D5, D6, D7) are presented as RDM (Raw Device Map – individual abstracted device) to different VMs (VM1, VM2, VM3). This allows the guest OS to format the file system directly on the raw device file. Application and data on the VM are configured in a native file system of the guest OS created on the RDM.

While the VMs benefit from VMFS services, as per the application demand, the MPX200 solution provides many individual devices (Vdisks) to those VMs.

The MPX200 provides more ESX server connectivity and high performance via the 10G option and is an excellent solution for traditional ESX server implementations and deployments.

HP StorageWorks MPX200 Solution—improved implementation methods for VMware

This section describes the current technology compared to the improved methods with the EVA and MPX200. Previous methods dictated that the storage initiator, either hardware or software, reside at the hypervisor level. This is because most mid- to high-end storage arrays allowed 128 to 512 host connections. Each initiator on the server side consumed one host connection on the array side. Therefore, enough array host connections were not available for every virtual machine on every physical platform. For example, one hypervisor can support up to 35 hosts. Typical cases have 8 to 12 virtual machines on one hypervisor. A mid-sized array would run out of host connections after, as few as, 11 hypervisors were deployed.

Even if enough host connections were available from the array, the IT manager would still run into the limit of 256 LUNs within the storage area. Every hypervisor needs to be exposed to every LUN within the environment for VMotion to work. This inherent limitation has a workaround within the hypervisor. If more than 256 virtual machines are deployed, the respective hypervisors carve the assigned LUNs to provide partial LUN access to each guest OS. This is counterproductive to the existing policies for security, regulatory compliance, and streamlined storage management.

This results in a circular paradox: you want unique LUNs assigned to each guest OS, but you cannot have LUNs independent to a guest OS unless it is running its own initiator, and you cannot run more initiators without more array host connections, and without more host connections, you cannot assign unique LUNs to each guest OS.

The breakthrough initiator virtualization technology to overcome this paradox is in the MPX200. By bridging the Fibre Channel and iSCSI networks, the MPX200 can provide host connectivity to the EVA storage system but the solution goes well beyond the Fibre Channel-iSCSI routing capabilities. The MPX200 uses (at most) four host connections in its maximum configuration. However, those four connections can be virtualized to over 1024 iSCSI initiators. The paradox exists from both sides: array connections are preserved and iSCSI initiators are now available for each guest OS. The MPX200 iSCSI solution overcomes the traditional VMware implementation limitations in the following two possible implementations:

Implementation Option 1

Architecture

The goal of this architecture is to provide performance, flexibility, and scalability in storage configuration in a virtual server environment by combining access to storage LUNs by using both iSCSI initiators configured in ESX server and on individual guest operating systems.

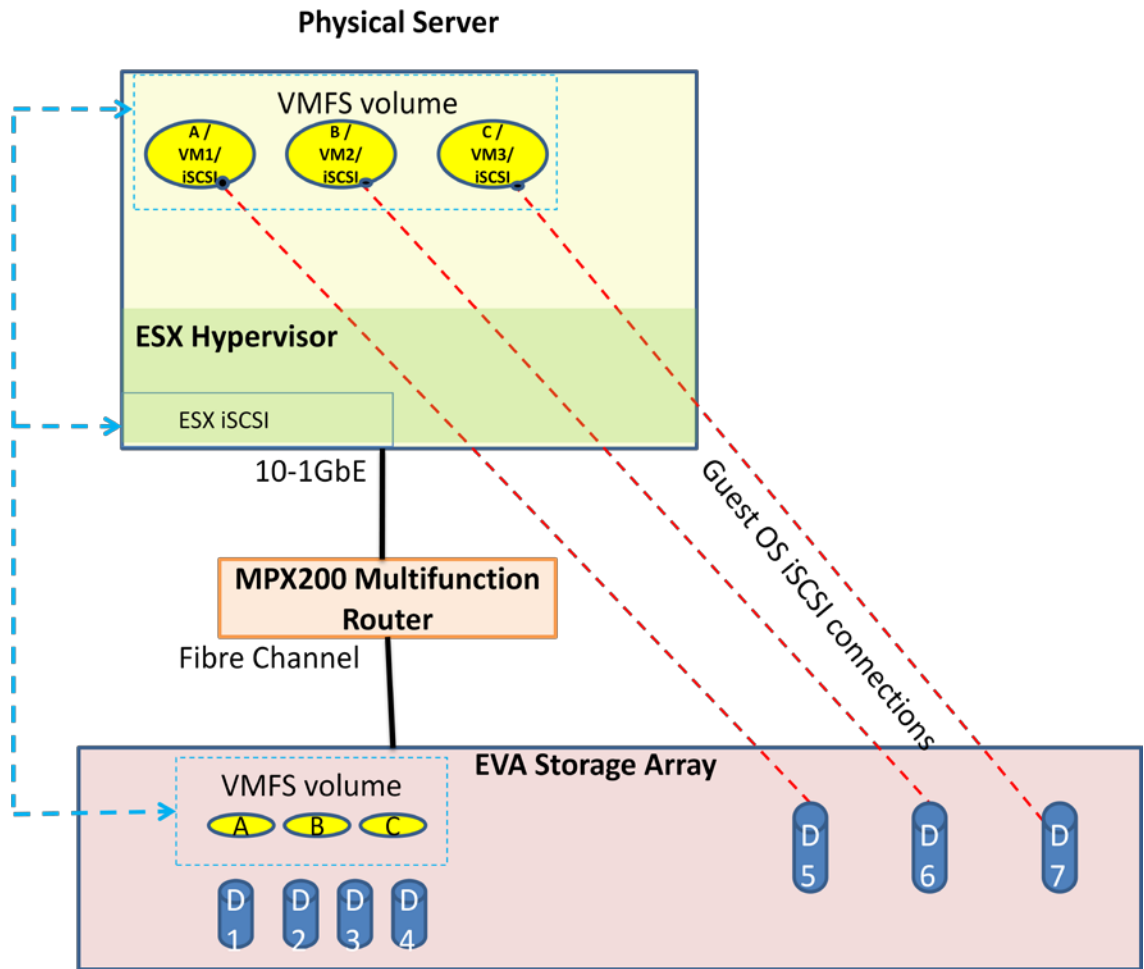
The VM is configured in the VMFS volume using ESX server's iSCSI initiator. The guest OS is installed on this volume. The data LUN is configured using the iSCSI software initiator at the VM's guest OS. The software iSCSI initiator in the guest OS accesses the MPX200 iSCSI solution for EVA and Vdisks in it. The Vdisks are presented to individual guest operating systems. The architecture facilitates that the guest OS owns and manages the LUNs. The VM, in the virtual environment, is like a standalone system running its own software iSCSI.

Benefits

As a result, the guest OS will have all its features available to the devices. In the case of EVA, customers can continue to use their existing business compliance software licenses and policies. In addition, Microsoft's multipath I/O (MPIO) for multipathing and MSCS for clusters can also be deployed at the guest OS level. Hence, the VM is managed as a standalone system while it enjoys the benefits of being in virtualized environment. This also is very useful for customers moving from a standalone, physical server environment to a virtual server environment.

Figure 2 shows the improved way of configuring iSCSI in a hypervisor and on virtual machines.

Figure 2: Implementation Option 1



The physical EVA Vdisks volumes are pooled into a single volume extent and the VMFS file system is created. The number of Vdisks in the pool can be one Vdisk or many Vdisks. Because MPX200 supports many Vdisks, it provides more options to pool Vdisks and present them to the ESX servers.

- You can choose a large single Vdisk and configure VMFS and share it across many ESX servers. This option simplifies the VMFS and Vdisk management.
- You can choose more than one Vdisk in the pool and share the VMFS volume across all the ESX servers. This option also simplifies the VMFS and Vdisks management. We will explain this in more detail in option 2.
- You can choose one Vdisk in the pool and dedicate it for one ESX server. Every ESX server will have its own Vdisk and VMFS volume on them. As the resources are not shared by ESX servers, there will be some improved performances that some applications can take advantage of. However, the management of VMFS and Vdisks are highly involved.
- You can choose more than one Vdisk in the pool and dedicate it for one ESX server. Every ESX server will have its own Vdisks pool and VMFS volume on them. As the resources are not shared by

ESX servers, there will be some improved performances that some applications can take advantage of. However, the management of VMFS and Vdisks are highly involved.

In [Figure 2](#), Vdisks D1, D2, D3, and D4 are pooled together and a VMFS file system volume is created. In the VMFS volume, Virtual Machine Disks (vmdk) are carved out for every VM. As the figure shows, vmdk A, B, and C are carved out for VMs. VM1 is installed in vmdk A, VM2 in vmdk B, and VM3 in vmdk C. Microsoft Windows guest OS is installed in every VM. Microsoft's software iSCSI initiator is configured in every VM. The software iSCSI initiator in the VM is configured to connect to MPX200. Vdisk D5 is presented to the iSCSI initiator in VM1, Vdisk D6 is presented to the iSCSI initiator in VM2, and Vdisk D7 is presented to the iSCSI initiator in VM3. Because the VM has direct ownership of its LUNs, it has the flexibility of managing those LUNs it wants to. The iSCSI EVA Vdisks in the guest OS can be configured with the native file system of the guest OS, and the application and data can be stored there.

Implementation Option 2

Architecture

This architecture provides a high performance solution in a virtual server environment. The increase in performance is because the Raw Device Mapped (RDM) volume is configured in the MPX200 iSCSI SAN and a single Vdisk is dedicated for a single VM. A single guest OS is configured in that dedicated RDM volume. For the data LUNs, the iSCSI software initiator is configured in the guest OS's. The Vdisks are presented to individual VMs. The architecture facilitates that the guest OS's owns and manages the LUNs. The VM, in the virtual environment, is like a standalone system running its own software iSCSI.

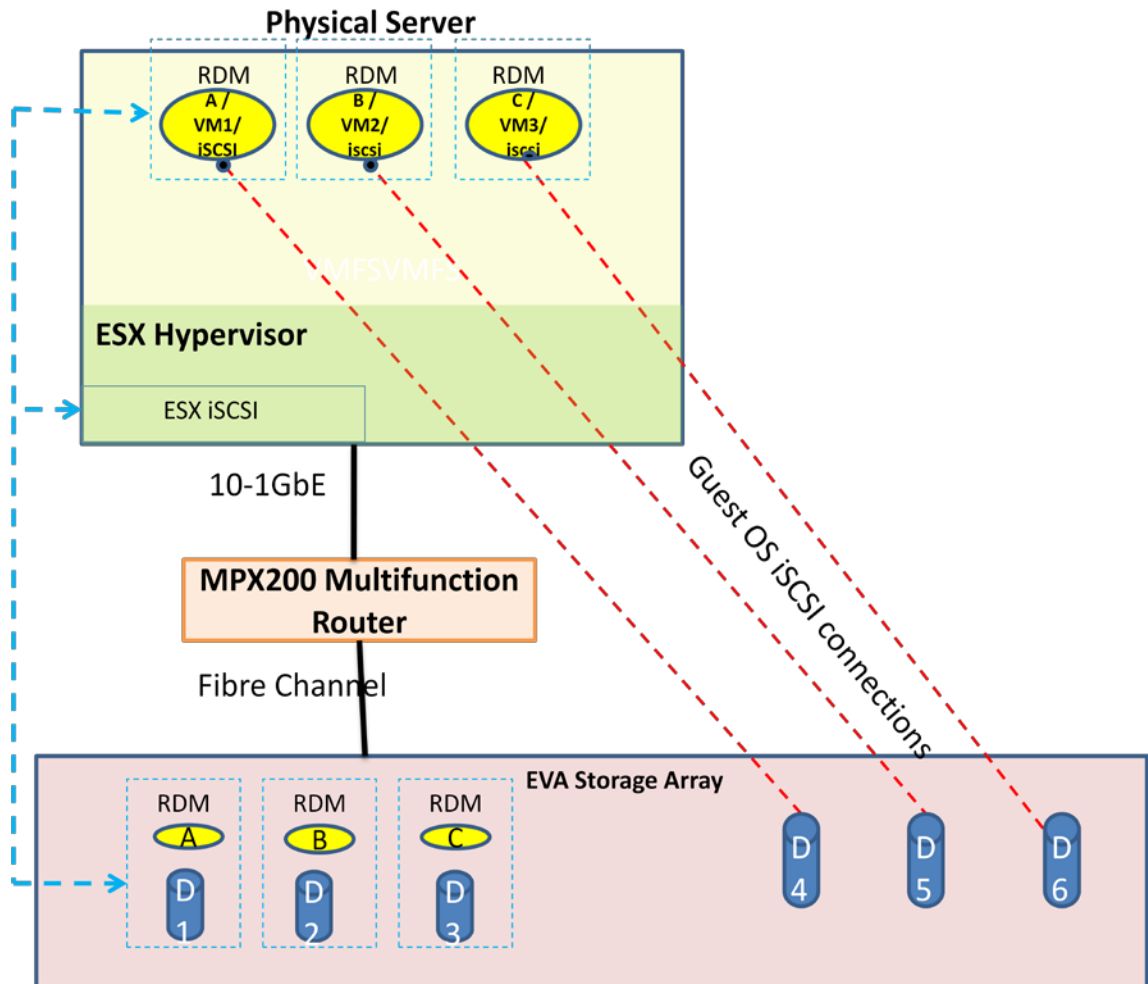
Benefits

As a result, the guest OS will have all its features available to the devices as in implementation Option 1. In addition, the VMs are configured entirely on the SAN, thereby enabling the benefits of VMware's traditional implementation. Because a single VM runs on a single RDM LUN, this reduces the number of shared resources; leading to high performance VMs. This also leads to simplification of VM management because a larger number of disks are supported in the MPX200 EVA iSCSI solution to allow individual disks for every VM.

This option allows a maximum of 512 high-performing virtual machines that own one iSCSI LUN, each in a single MPX200 configuration [512 for installing the guest OS and 512 data luns].

[Figure 3](#) shows the improved way of configuring iSCSI for high performance virtual machines in a hypervisor.

Figure 3: Implementation Option 2



The physical EVA Vdisks volume is configured as the RDM volume. In this option, only one Vdisk is used to create the RDM volume. Thus, RDM A is created on Vdisk D1, RDM B is created on Vdisk D2, and RDM C is created on Vdisk D3. VM1 is installed in RDM A, VM2 in RDM B, and VM3 in RDM C. Microsoft Windows guest OS is installed in every VM. Microsoft's software iSCSI initiator is configured in every VM. The software iSCSI initiator in the VM is configured to connect to MPX200. The Vdisk D4 is presented to the iSCSI initiator in VM1, Vdisk D5 is presented to the iSCSI initiator in VM2, Vdisk D6 is presented to the iSCSI initiator in VM3. Because the VM has direct ownership of its LUNs, it has the flexibility of managing those LUNs it wants. The iSCSI EVA Vdisks in the guest OS can be configured with the native file system of the guest OS, and the application and data can be stored there.

The basic configuration would have one LUN defined for each guest OS that is managed by the hypervisor and contains the VM's boot image. The hypervisor would also run its own iSCSI initiator. As the VM is installed in a single Vdisk RDM volume, the possibility of any VMs sharing it is eliminated; hence, this configuration is desirable for high-performance applications, while the VMs benefit from owning and managing its LUNs.

HP StorageWorks MPX200 Benefits

In a virtual server environment, this means that each guest OS can have its own initiator without relying on the hypervisor to secure and manage data LUNs. Storage can be allocated and expanded

based on the requirement of each guest OS or application instance. This frees the IT Manager from the constraints of the 256 LUN limit from the following perspectives:

- The hypervisor is no longer a bottleneck for management, growth, and data security.
- Array connections are not consumed by the virtual machines.

With architectural support for over 1,024 iSCSI initiators per MPX200, IT staff can continue deployments of virtual machines, without changing the storage infrastructure, its management, or security policies.

This enables powerful solutions for the storage manager. Virtual machines can be deployed without having to over-allocate data. New data LUNs can be easily assigned any time in the future without considering LUN limits or hypervisor intervention.

As the customers moving towards Virtual Desktop environment and implementing VMware's Virtual Desktop Infrastructure (VDI) to benefit out of Virtual desktop paradigm, all the implementation options are applicable to VDI as well. The customer can choose to install either Windows servers or Desktop Windows in all the implementation options discussed in this paper.

Conclusion

The MPX200 10-1 GbE blade allows high performance virtual machine deployment with 10GbE iSCSI connectivity. This blade supports up to 600 iSCSI initiator connections to the EVA with architectural support of up to 1,024 iSCSI initiators. As a result, each virtual server can be configured with its own iSCSI initiator and securely manage its own LUNs. This ensures that management policies can be seamlessly maintained throughout the VM deployment. Configuring iSCSI initiators on VMs also enables administrators to overcome the hypervisor limitation of 256 LUNs.

For more information

- For more information about the product, see: <http://www.hp.com/go/mpx200>
- For more information about configuration, see the HP StorageWorks SAN Design Reference Guide available at: <http://www.hp.com/go/sandesignguide>

Technology for better business outcomes

© Copyright 2010 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

4AA1-6606ENW, June 2010



Get connected
www.hp.com/go/getconnected
Current HP drivers, support & security alerts
delivered directly to your desktop

