ESG Lab Review™

Emulex Optimized Server Virtualization

A Product Capabilities Review
by
ESG Lab
May 2007

Authors:
Brian Garrett, Mark Bowker, Tony Palmer

Table of Contents

ESG Validation Reviews

The goal of the ESG Lab reports is to educate customers about specific storage-related products including storage systems, backup-to-disk solutions, storage management applications, backup/recovery software, storage virtualization platforms, etc. ESG Lab reports are not meant to replace the necessary evaluation process that end user customers should conduct. ESG Lab reports are designed to provide insight to what is compelling about various products and how they can solve customer problems. ESG Lab provides third-party expert perspective based on ESG Lab analysis and interviews with customers using these products in production environments.

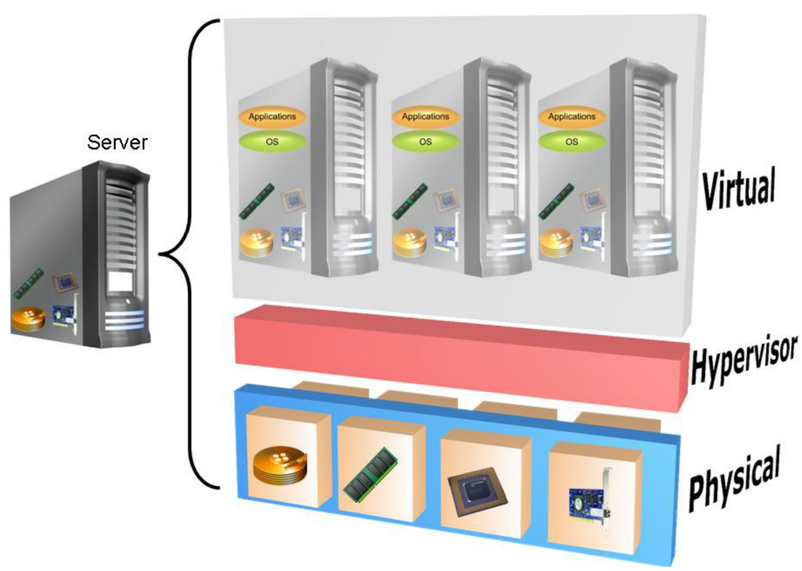
Introduction

A growing number of organizations are embracing virtual server technology to reduce server and support costs and enhance disaster recovery capabilities. Leading systems, storage and software vendors have been working to eliminate the inhibitors to virtual server deployment in the data center. Emulex in particular has been spearheading storage connectivity enhancements designed to meet the security, scalability and support requirements of virtualized data centers. For example, Emulex has taken a leading role in the development and adoption of N_Port ID Virtualization (NPIV).

N_Port ID Virtualization (NPIV) is an industry standard optimized for mixed virtualized server workloads. Emulex and IBM jointly invented NPIV, and were first to market with NPIV support for use in the demanding IBM zSeries mainframe product line. Emulex is working with server virtualization vendors including VMware, Microsoft and Xen, with general availability beginning in 2007. In this ESG Lab Review, we will introduce the role of NPIV and explore the architectural advantages of the Emulex LightPulse® architecture that make it an ideal solution for use in virtualized server environments.

Server virtualization is used by IT professionals to consolidate multiple applications onto fewer servers. As shown in Figure One, a software layer, generically referred to as a Hypervisor, is run on a physical server to divide the server into multiple Virtual machines. VMware ESX is an example of an industry leading product that uses Hypervisor technology. The Hypervisor is the software layer that performs the virtualization magic – turning one set of physical storage, memory, processors and network connections into multiple sets of resources – one for each virtual machine. As a result, multiple applications and operating systems can be deployed on the same physical machine.

Figure One: Server Virtualization



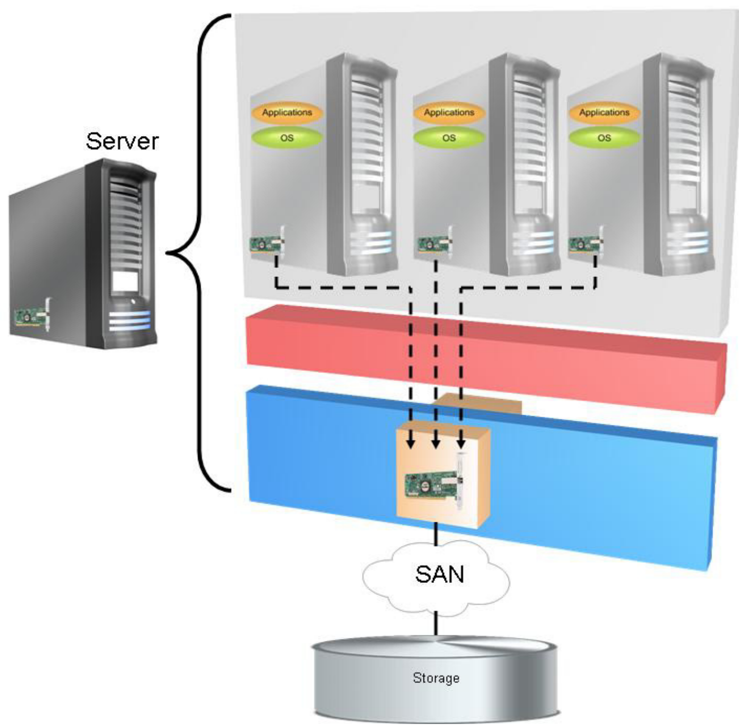
Why This Matters

Server virtualization enables server consolidation. Server consolidation reduces server acquisition and administration costs while increasing server utilization and enhancing application portability and availability.

Virtual Servers and Shared Storage

Virtual servers are usually connected to a shared pool of storage, as shown in Figure Two. The applications and operating systems on the virtual servers are using a Fibre Channel host bus adapter to access a disk array connected to a Storage Area Network (SAN). Instead of relying on the hard drives inside a server, an HBA is used to access SAN attached storage within a shared disk array. This level of shared access is needed to take advantage of the high-availability and workload management features of hypervisors including VMware's VMotion.

Figure Two: Multiple Virtual Machines sharing a single Physical HBA

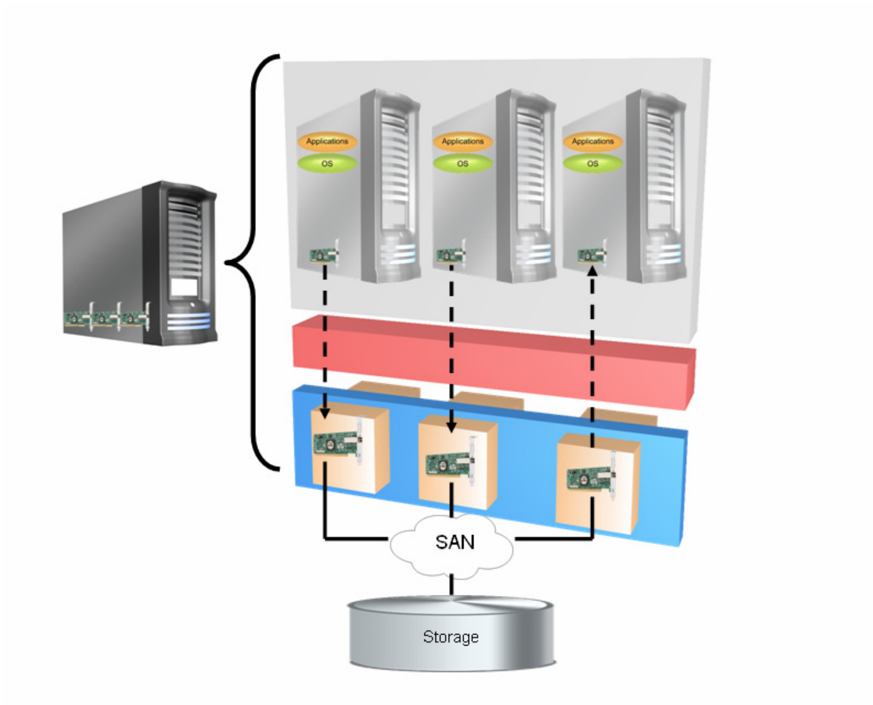


Why This Matters

SAN attached shared storage, as a foundation for virtual server architectures, reduces storage administration and backup costs. SAN attached storage has even greater benefits when it is used to eliminate internal hard drives within virtualized high density and bladed servers – power and cooling expenses are reduced, storage utilization is increased and the downtime associated with server upgrades is reduced.

Shared storage in a virtual server environment clearly has compelling benefits. What may not be as clear are the subtle implications that this has on the way shared storage is deployed and managed. In particular, how can IT administrators ensure that virtual operating systems and applications do not overwrite data belonging to another virtual machine? For example, imagine test and production versions of a database application deployed on a pair of virtual machines running within the same physical server. How can an IT administrator make sure that developers testing new changes do not have access to production data? One option is to give each virtual machine its own physical host bus adapter, as shown in Figure Three. In this case, each of three virtual servers is using its own physical Host Bus Adapter within a server.

Figure Three: A dedicated Host Bus Adapter for each Virtual Server



Dedicating HBAs to each virtual machine is pricey and does not deliver much additional value for the investment. While you get hardware isolation and the ability to use zoning and LUN masking, you incur the cost of an HBA for each virtual server (two, if high availability is required) and the additional switch ports needed to connect each physical HBA port to your SAN infrastructure.

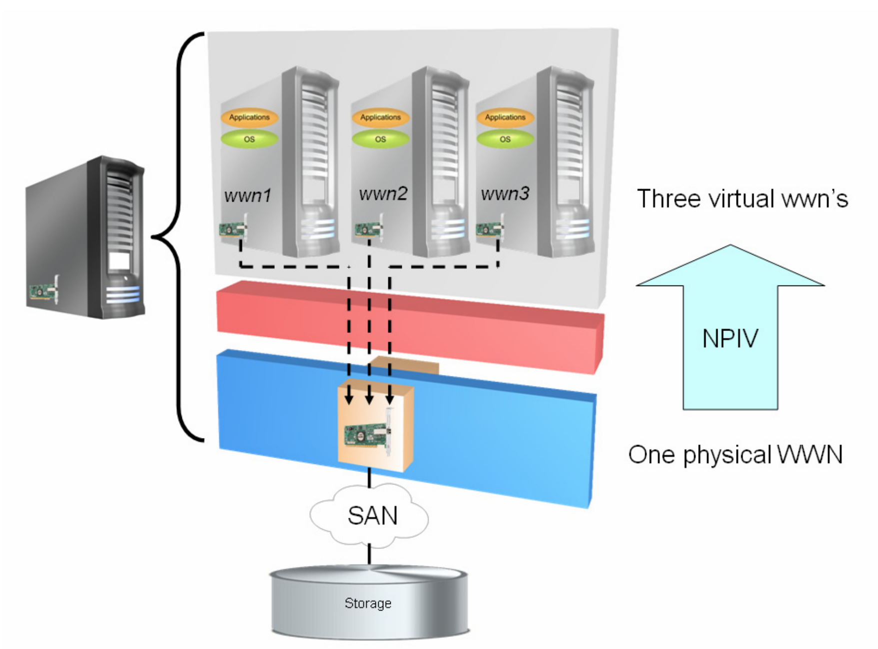
At first blush it might seem that a physical host bus adapter for each virtual machine is overkill. Why not use the partitioning capabilities of the operating system or the virtualization software to slice a pool of storage accessed through a single host bus adapter? IT administrators are wise to steer clear of this option because it reduces the level of security and isolation compared to traditional zoning and LUN masking implemented within the Fibre Channel infrastructure. Just as important, it changes the way that shared Fibre Channel storage is provisioned, secured and managed.

N_Port ID Virtualization

The problems that NPIV was born to solve are similar to those addressed with zoning and access control when Fibre Channel was first invented. In a traditional SAN environment, zoning and LUN masking (also known as access control) are used to ensure that a particular system can only see and update the storage that belongs to it. Zoning facilitates and restricts communication between ports in a SAN Fabric. LUN masking allows a storage administrator to restrict access for specific volumes. Taken together, these techniques ensure that a host accessing storage in a SAN has exclusive access to its volumes. This has proven especially valuable in heterogeneous environments where Operating Systems (Windows, for example) tend to write a signature on every disk it can see, making volumes within the same storage array unusable by any other operating system. Besides this basic need to avoid one operating system from clobbering another, the same techniques are used by data center managers to enhance and ensure secure isolation of information assets between disparate users, applications and groups.

In a virtualized environment, these requirements still exist. In fact, they become critical when multiple virtual machines are accessing the SAN through the same host bus adapter. NPIV is the industry-standard technology that makes this all work. NPIV is defined in the Fibre Channel Framing and Signaling standard as a facility for sharing a single physical port among multiple virtual ports. NPIV provides a method by which each virtual server presents a unique World Wide Name (WWN) address to the physical SAN. This gives multiple virtual servers secure access to a shared pool of SAN attached storage using the same physical port. For example, Figure Four shows how the NPIV layer implemented in a host bus adapter and a driver from Emulex can be used to present multiple virtual world wide names – one for each virtual server, using a host bus adapter with only one physical world wide name.

Figure Four: N_Port ID Virtualization (NPIV)



Why This Matters

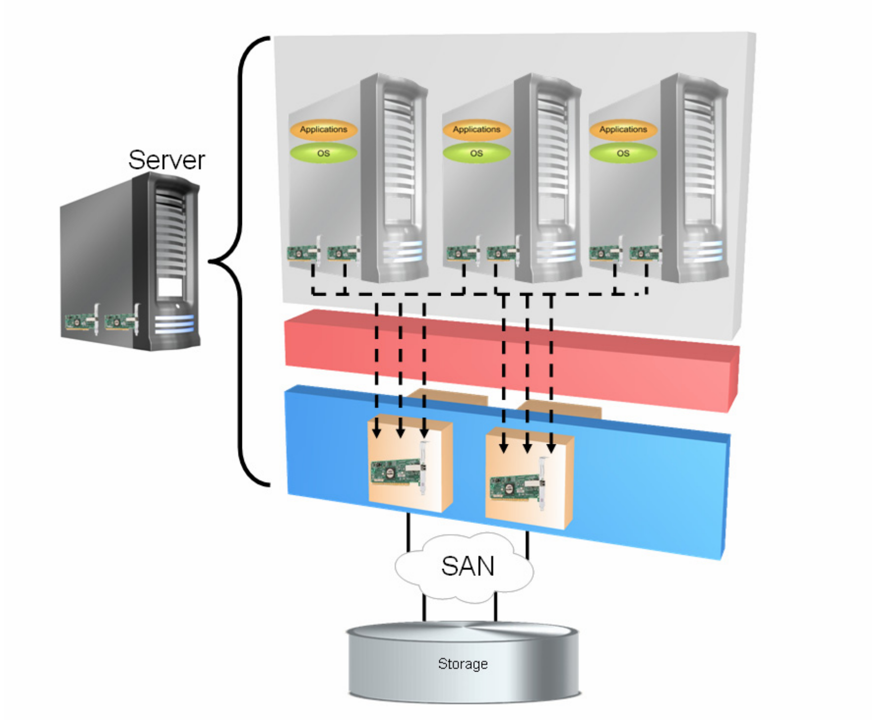
Using NPIV, zoning and access control can operate on each virtual machine independently without incurring the cost of a dedicated physical HBA port for each virtual server. This gives each virtual machine secure access to its required storage resources using the same tools and techniques familiar to the SAN administrator and allows for traffic identification and Quality of Service administration at the SAN layer. Additionally, NPIV technology allows end users to maintain the same SAN Best Practices on virtualized servers as on traditional hardware-based servers.

Enabling High Availability

NPIV reduces the cost and complexity of consolidating clustered, highly available applications that rely on access to shared pool of storage. For example, consider the common use case where development, test and production have been deployed on virtual machines within the same physical server as shown in Figure Five. Multiple virtual ports have been created on physically redundant HBAs to provide a secure and highly available storage path for each virtual machine. Traditional clustering software, which is accustomed to accessing redundant paths into an FC SAN, works without modification. Using NPIV, multiple virtual machines can now share the same physically redundant SAN connection and IT administrators do not have to worry about production data corruption due to changes in development and tests.

In addition, NPIV enables selective fabric and array configuration at the virtual machine level, enabling the SAN administrator to guarantee fabric bandwidth (using fabric switch tools) or customize storage attributes (array tier, mirroring, frequency of backup) to the applications hosted on specific virtual machines.

Figure Five: Clustered Virtual Machines Sharing a Highly Available SAN Connection



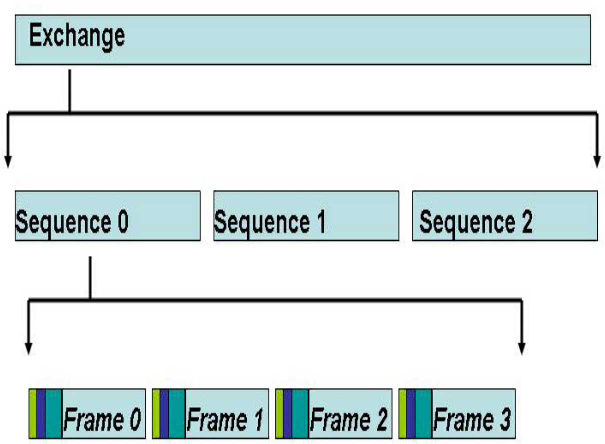
The virtual WWN addresses presented using NPIV are 'portable,' meaning that they are not tied to a specific physical HBA. This enables seamless migration of virtual machines across hardware platforms. The WWN is associated to a specific virtual machine and when the virtual server is moved to a new physical server, the WWN moves with it.

Performance Optimized

Emulex host bus adapters are optimized for mixed workload performance for the majority of virtual servers using SANs today. Mixed workload performance is especially crucial in consolidated, virtual environments where multiple virtual servers and applications are all accessing storage through the same set of physical HBAs. ESG Lab has reviewed the excellent mixed workload performance characteristics that Emulex provides with their capability of interleaving Fibre Channel data transfers at the frame level.

To illustrate why frame-level multiplexing has such an impact, let's start with the basics of a Fibre Channel communication exchange. An I/O transaction in Fibre Channel is called an Exchange. Exchanges contain one or more sequences, which in turn contain one or more frames, as shown in Figure Six. Frames can be 512, 1024 or 2048 bytes in length, but 2048 is used almost universally. Think of the frame as a word, the sequence as a phrase and the Exchange as an entire conversation.

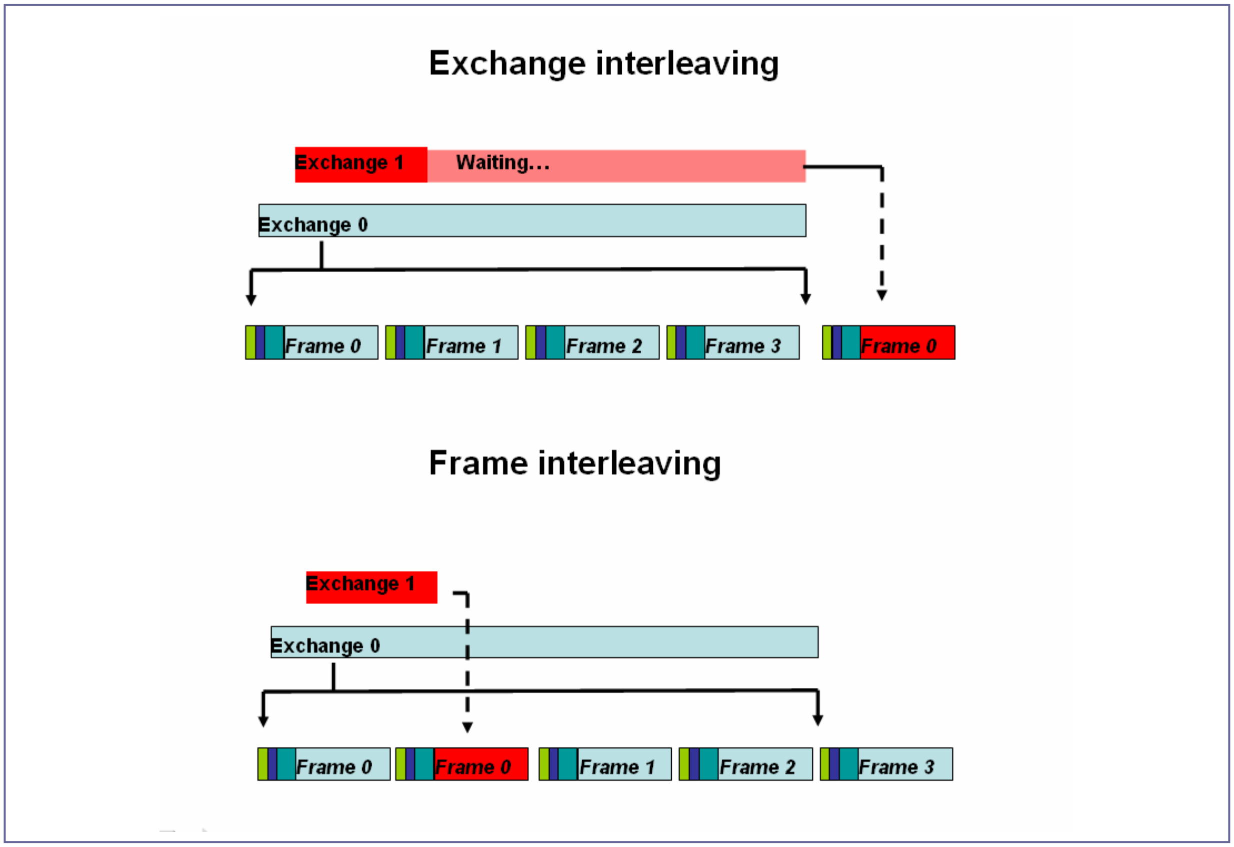
Figure Six: The Anatomy of a Fibre Channel I/O Exchange



Frame interleaving allows transfers to be inserted between the frames of another sequence instead of having to wait for the end of the conversation. The end result is much like the difference between having a conversation on a phone versus a walkie talkie. With the walkie talkie, you have to wait until the other party finishes speaking before you can begin your transmission. Frame-level multiplexing is more like a phone since you can begin talking whenever you are ready.

The difference between Exchange and frame-level multiplexing is illustrated in Figure Seven. A data transfer conversation begins on the far left (Exchange 0). That conversation is broken into four frames (frames 0, 1, 2 and 3). A second conversation (Exchange 1) begins shortly after the first conversation has begun. When the traditional exchange interleaving method is used, the first frame of the second conversation can not be transferred until the first conversation (Exchange 0) is complete. With Emulex frame level interleaving, the second conversation (Exchange 1) begins earlier and is interleaved with the first conversation. As a result, the second conversation begins transferring data and completes sooner. This translates into more efficient, reliable data transfer and improved performance.

Figure Seven: Exchange Vs. Frame Interleaving



The advantages of frame-level multiplexing are most pronounced when conversations with large payloads (e.g. backup traffic composed of large sequential writes) are intermixed with performance sensitive transactional workloads (e.g. interactive database applications composed of smaller reads and writes).

Why This Matters

Mixed workloads are common in a virtualized server environment where multiple applications are running over the same host bus adapter. Emulex host bus adapters with frame-level multiplexing are optimized for mixed workloads.

ESG Lab View

Server virtualization technology has matured in recent years and is being adopted by a growing number of IT managers looking to reduce hardware and management costs through server consolidation. This trend goes hand in hand with the increasing adoption of storage consolidation. IT administrators are reducing hardware and management costs as they centralize the management of a consolidated pool of Fibre Channel attached storage. When you put the two together, server and storage consolidation can be used to create a centrally managed platform for the cost effective deployment of grid enabled applications.

NPIV increases the security of virtual servers by enabling secure access to shared Fibre Channel storage using the zoning and LUN masking techniques familiar to the SAN administrators. NPIV also reduces cost and complexity. Emulex was the first to market with industry standard NPIV support for use in the demanding IBM zSeries mainframe environments. All of the major storage switch vendors support NPIV and NPIV enabled switches from Brocade and Cisco have been tested with Emulex virtual HBA technology to ensure performance and compatibility. As server virtualization vendors, including VMware, Microsoft and Xen, bring NPIV support to market in cooperation with Emulex, ESG Lab believes that IT managers formulating server and storage consolidation plans for 2007 and beyond should consider Emulex as the keystone of an emerging grid-enabled architecture.

With industry leading NPIV support and an architecture optimized for mixed virtualized server workloads running on servers equipped with the latest 64 bit Intel Xeon and AMD Opteron processors, ESG Lab believes that Emulex host bus adapters are ideally suited for applications being deployed on a virtual server based infrastructure.