White Paper

Optimizing the Virtual Data Center with Data Path Pools

EMC PowerPath/VE

By Bob Laliberte

February, 2011

This ESG White Paper was commissioned by EMC and is distributed under license from ESG.

Contents

Expanding Infrastructure Virtualization

Limited budgets and underutilized data center resources are leading many organizations to implement virtualization technologies for greater efficiency. In fact, according to ESG's 2011 IT Spending Intentions Survey, increasing server virtualization remains the number one priority among mid-sized and enterprise organizations—as it was in 2010.[1] Diving a little deeper, respondents were asked what business initiatives they expected to have the greatest impact on IT spending decisions in the next 12 to 18 months. As Figure 1 shows, initiatives aimed at cost reduction and business process improvement remain the top two initiatives. This is in sync with the server virtualization revolution, which is designed to reduce both CAPEX and OPEX; server virtualization requires fewer physical resources and maintenance and licensing costs are generally lower. It also delivers business process improvements such as faster provisioning, higher application availability, and better backup and recovery.

Figure 1. Business Initiatives with the Greatest Impact on IT Spending Decisions

Source: Enterprise Strategy Group, 2011.

It should be noted that the gap between cost reduction and business process improvement priorities is diminishing. The importance of cost reduction dropped from 54% to 42% of respondents since 2009 while business process improvement importance has increased from 31% to 33%. This indicates that organizations are becoming more focused on improvements such as virtualization and automation that increase efficiency.

There are other drivers of virtualization as well. For example, software costs are often lower with virtual servers than with multiple physical servers in the cloud. Server virtualization standards are speeding implementations and

[1] Source: ESG Research Report, *2011 IT Spending Intentions Survey*, January 2011.

support for multiple operating systems on a single hypervisor platform reduces complexity. As a result, organizations are experiencing increased application uptime, simpler management, reduced management costs, and better disaster preparedness.

In addition, recent ESG research focused on server virtualization noted an interesting trend: the more experience organizations have with server virtualization, the greater their benefits.[2] That is, organizations that have only just begun to implement server virtualization (measured by the scope of deployment, production virtual machine (VM) ratio, efficiency, and workload penetration) experience fewer benefits than those organizations with extensive virtualization deployments. This is good news since all signs point to massive server virtualization growth over the next two years: ESG respondents reported expectations of increasing their server virtualization deployments, consolidating more and more virtual machines on each physical server, and deploying more mission-critical, production workloads on VMs.

What will it take to build a data center optimized for all this? Most industry experts agree that a comprehensive approach incorporating the server, network, and storage domains is the answer. Most begin by leveraging server virtualization to create a pool of server resources that can be distributed and shared and optimizing storage by creating pools of capacity that can also be shared. But these virtualized workloads usually result in much more consolidated IO from a single physical server and its data paths. To fully optimize the data center, the network data paths must also become a resource pool (via multipathing). Without that, IO can become a bottleneck between the newly efficient server and storage pools.

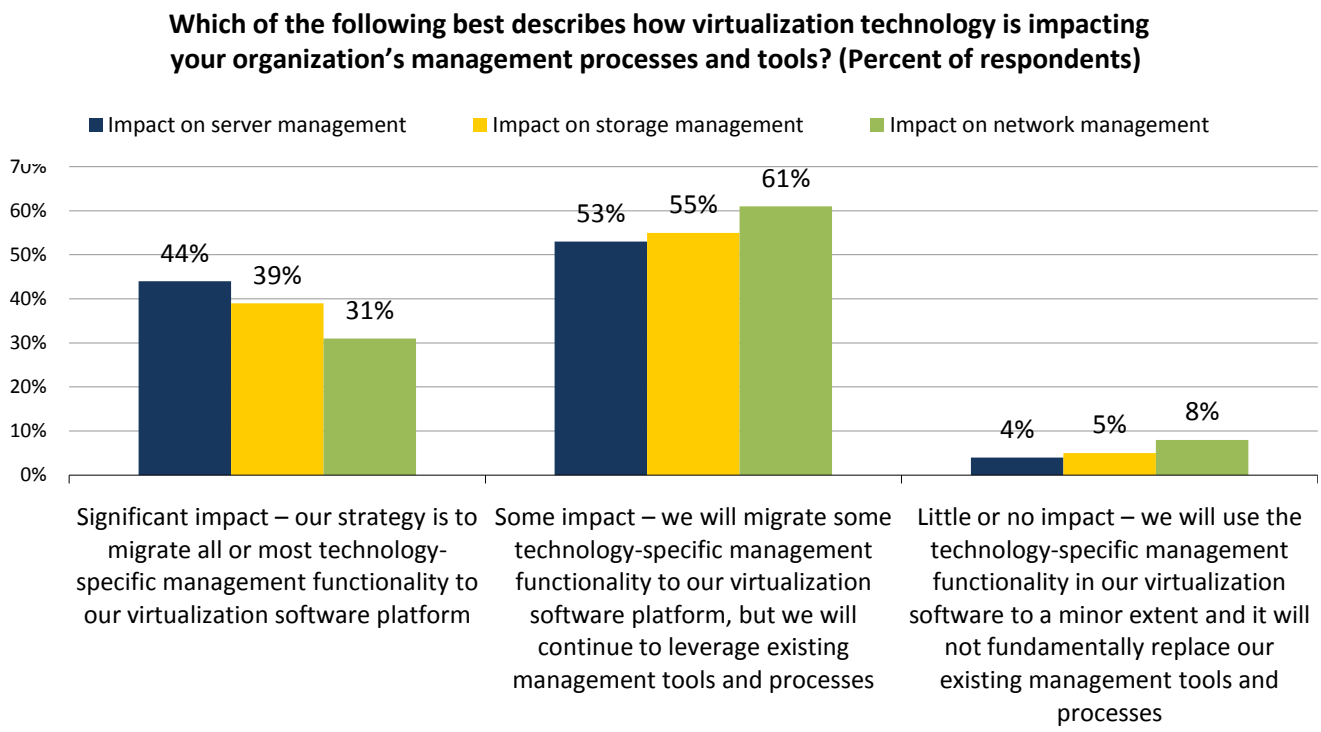
Virtualization Comes with Hurdles

Creating a fully optimized and dynamic data center is not a simple task; several challenges must be overcome in order to realize the benefits. These challenges span all domains and include:

- **The impact of virtualization on management tools and process.** Organizations should be aware of the impact creating these resource pools will have on existing management tools and processes. As Figure 2 demonstrates, in a recent server virtualization survey, the vast majority of respondents reported some impact to significant impact on server, storage, and network management and less than 10% indicated that virtualization was having little or no impact. As server virtualization environments continue to mature, organizations will virtualize more and more mission-critical production workloads and demand tighter integration and cross-functional, virtualization-aware management tools.
- **IO paths becoming bottleneck as consolidation ratios increase.** In dynamic virtual environments, it will be incumbent upon IT to get the most out of their data paths. As organizations continue to drive higher VM consolidation ratios on individual servers, IO can become a bottleneck. Complicating factors include pushing more IO into the data path than it can handle; constant, manual monitoring and tuning as IT attempts to optimize IO as virtual machines scale; and the inability to view IO end to end. Ultimately, IO problems impact application performance and availability—for which there is little tolerance now and even less as more production workloads are virtualized. It should be noted that virtualization users expect to dramatically increase the consolidation ratio of VMs to physical servers in the next two years. While nearly three-quarters of organizations have achieved consolidation ratios of at least 5:1, nearly one-third expect that in two years it will be 25:1. As virtual environments grow into this hyper-consolidation, problems occur. Interruptions can be created by IO-intensive applications and it becomes extremely difficult to manually map the environment to ensure sufficient IO response times.

[2] Source: ESG Research Report, *The Evolution of Server Virtualization*, November 2010. All research references come from this report unless otherwise stated.

Figure 2. The Impact of Virtualization Technology on Management Processes and Tools



Source: Enterprise Strategy Group, 2011.

- **Maintaining availability and performance in dynamic environments.** Maximizing application uptime to satisfy both management and users can also be challenging in highly virtualized dynamic environments. Inefficient, manual processes are not sufficient for quickly balancing IO loads. Dynamic environments with moving pieces make it difficult to see where IO is needed and how to get it there; features such as VMware vMotion, High Availability, Dynamic Resource Scheduler, and vStorage APIs for Array Integration (VAAI) all contribute to complicating IO performance management. This is no small matter to users as noted by the fact that performance issues appear third on the list of factors preventing organizations from using virtualization technology more pervasively. If IT is going to deliver the performance levels demanded by the business, these subpar deployments will never fly. Similarly, IT must be able to ensure seamless IO path failover during array upgrades, unexpected path failures, and unplanned outages or disasters. Given the increased virtual machine density, mixed workloads, and production application support expected, performance and availability will be in greater demand—and more difficult to ensure.
- **Reducing the cost of creating a highly virtualized, dynamic environment.** ESG noted that "vendor licensing/pricing for virtual technologies" is another of the top factors preventing organizations from more pervasive use of virtualization. For those in multi-hypervisor environments, building standards around multipathing in shared SANs is an issue as well; without them, deployment and management is complex and unreliable. In addition, having to manage the physical environment with one set of tools and the virtual environment with another requires extra cost and effort and leads to inefficiency. Tools that can standardize on a single solution to manage both are highly valued.
- **Resolving data path problems in virtualized resource pools**. Path management configuration is also a key consideration: it is complex and expensive to manually configure paths with virtual machines coming and going, growing, and changing. Correcting path failures and poor configurations interrupts operations and takes effort and staff time. Part of the problem is that it is difficult in this multilayered, dynamic environment to have adequate visibility to find the causes of failures, making it hard to resolve them quickly and efficiently.

- **Lack of virtualization skills, tools, and insight.** Networking, server, and storage employees reported challenges that included not only the cost of virtualization infrastructure in each category, but also the lack of virtualization knowledge and skills. In addition, networking employees pointed out that existing network management tools are built for physical, not virtual, devices and they cite difficulty in sizing the bandwidth required to support the virtual server environment. Similarly, when asked about the impact of server virtualization on their technology areas, storage and network employees noted an increased need for collaborating with other infrastructure groups in addition to increased use of SAN-based storage and creation of more network traffic. All these data points combine to demonstrate the importance of end-to-end visibility across the virtualization stack. Tools that enhance visibility and enable IT to see what's going on from any technology perspective will help smooth the road to greater virtualization.

As mentioned previously, extending virtualization deployment seems to increase the benefits experienced, so tools and technologies that enable that expansion can clearly be advantageous. However, not all solutions are created equal; organizations should be advised that many of the "free" native solutions bundled with virtualization technology often require manual tuning and offer reduced monitoring, not to mention a new learning curve and support for only virtual environments. Instead, organizations should investigate well-proven solutions from the physical world that have been adapted to the virtual world.

Creating Pools of Data Paths with EMC PowerPath/VE

Based on reported challenges to expanded virtualization, what type of multipathing solution do users need? ESG research indicates that organizations are looking for:

- A single tool set to manage data path configuration, provisioning, and failover for both *physical* and *virtual* environments
- End-to-end IO visibility across the virtual environment
- Automation of multipathing and IO load balancing
- Seamless path failover and recovery
- Improved performance and path optimization
- Lower costs, regardless of scale

Organizations are looking for greater efficiency, lower costs, the ability to maximize IO channel usage, and growth without worrying about data flows. The more automation and intelligence a tool includes, the more IT administrators are relieved of constant monitoring and adjusting data path usage for performance and availability. It is noted that virtualization is creating a shift from individual, silo-based administrative functions toward IT generalists that manage across the stack. Comprehensive tools with end-to-end visibility that deliver automation and intelligence will help make this possible.

With PowerPath/VE, users get a tool set to address the management challenges of dynamic virtual environments using VMware vSphere as well as Microsoft Hyper-V. Built on the gold-standard EMC PowerPath Multipathing for physical environments, PowerPath/VE also protects heterogeneous servers and storage, including non-EMC arrays. PowerPath has stood as the benchmark for path management solutions for a decade; with more than 30,000 unique PowerPath customers and one million host servers, it is a proven solution. It may not be EMC's most glamorous product, but it is surely one of its most respected.

As noted previously, server virtualization creates a shared pool of server resources and storage optimization creates shared pools of storage capacity to support applications. PowerPath/VE provides the dynamic multipathing required to create pools of data paths in the same paradigm. With PowerPath/VE, organizations can build fully optimized data centers based on shared pools of resources from servers through data paths to storage. PowerPath/VE increases the number of always active initiators, resulting in multiple data streams per host. IT can ignore the complexity of the virtual environment and simply assign all devices to all paths; PowerPath/VE automatically optimizes IO performance. Because PowerPath/VE resides in the hypervisor—below applications and

guest operating systems and above HBAs and storage—it offers heterogeneous support of guest operating systems and storage arrays. This eliminates the need to qualify VMs to path management products. PowerPath/VE serves as an infrastructure control point and works with any raw storage devices, volume managers, file systems, and applications. It offers a similar user interface across multiple operating systems as well as integrated monitoring via VMware vCenter Server and the new remote PowerPath CLI.

The true value of PowerPath/VE is intelligent path management across the entire shared SAN. It combines multi-path IO capabilities, automatic load balancing, and path failover and recovery and, together with PowerPath, provides a single tool set that works across heterogeneous physical and virtual server and storage environments—one tool for the entire job. A common path management solution delivered across multiple operating systems creates predictability and stability in the SAN, including Fibre Channel, iSCSI, and FCoE environments. PowerPath/VE is designed to maximize application availability, optimize performance, and automate path management while reducing complexity.

PowerPath/VE's patented algorithms and features make it unique among data path management solutions. They include:

- Load-balancing policies optimized for EMC Symmetrix VMAX, VNX, and CLARiiON arrays; a proprietary algorithm uses SAN-awareness and intelligent routing to adapt to load changes in the SAN. PowerPath/VE ascertains how paths are being used and how heavy the load is and routes IO accordingly.
- A proprietary load-balancing algorithm for non-EMC arrays.
- Automatic, non-disruptive path failover and failback (recovery)—providing peace of mind when dealing with mission-critical applications.
- Path testing with sensitivity toward failing paths and the ability to automatically restore recovered paths. Proactive failure detection is a unique feature of PowerPath and enables it to reroute IO requests before a failure occurs. Decaying and failed paths are quickly identified and removed from service, and restored paths are automatically returned to service.
- Periodic testing to determine which data paths are active or inactive. Load-based path testing minimizes CPU cycle use as busy systems successfully completing IO require less path testing.
- Automatic setting of load-balancing policies for path optimization.
- End-to-end SAN awareness for better visibility.

PowerPath/VE's intelligence provides a higher level of path management services. For example, when routing data paths, PowerPath/VE takes into account not just the number of IOs, but IO size and type, and the queuing delay on each path. Similarly, for failover and failback or recovery, PowerPath/VE considers the current workload, the load-balancing policy, and what paths are available and chooses the next best path after failure.

The Importance of Multipathing in Dynamic Data Centers

The features mentioned in the preceding section are particularly important when viewed in the context of expanded server virtualization. ESG asked respondents in various technology segments (servers, storage, networking, applications, and security) what developments they felt would enable more widespread usage of server virtualization. A common theme among these staffers was the need for better visibility end to end. They mentioned the need to simplify and consolidate data center switching; more features in virtual switches; greater virtualization training; and better integration between server, storage, networking, and virtualization technologies. Automated data path management with PowerPath/VE supports all these needs. It is interesting to note that a significant percentage of employees in all technology segments are seeking additional training to expand server virtualization; it may be that comprehensive tools such as PowerPath/VE, by automating tasks and optimizing resources, could minimize the amount of additional training required.

Importantly, PowerPath/VE's architecture supports multi-tenancy deployments for private and public clouds. Available bandwidth can be segmented on an application basis to ensure sufficient performance for each tenant. Specific applications can be assigned to dedicated channels and some applications can leverage more active paths as necessary. This should make PowerPath/VE attractive to service providers.

Finally, PowerPath/VE has been extensively tested by EMC E-Lab in hundreds of scenarios across multiple hardware and OS platforms using various protocols. It has been tested successfully in deployments using clusters, thin provisioning, VM partition mobility, various fault insertions, non-disruptive upgrades, SAN boot, and many more. As one of EMC's most widely tested products, its performance in all these scenarios is well established.

The Bigger Truth

An automated path management solution is essential for getting the most out of your virtual investment. ESG research indicates clearly that server virtualization deployments will expand dramatically over the next two years; a continuation of manual path management will not be able to sustain that growth. EMC PowerPath, including PowerPath/VE, delivers a single path management solution for physical and virtual servers, for EMC and non-EMC storage, and for the most popular operating systems.[3] Users gain virtual pools of path resources to accompany virtual pools of storage and server resources. PowerPath/VE leverages all IO channels for load-balancing and failover, providing predictable high performance. Its ability to constantly adjust IO path usage according to changes in IO loads from virtual machines ensures that applications retain the performance they need.

These features will be critical over the next two years as virtual machine density increases and as organizations virtualize more mixed workloads that include mission-critical applications. Instead of spending time managing IO and mapping bandwidth to applications and VMs, PowerPath/VE users can leverage its automation, intelligence, and end-to-end visibility to ensure performance, availability, and greater efficiency.

[3] PowerPath/VE supports VMware vSphere and Microsoft Hyper-V virtual environments. PowerPath Multipathing supports Windows, Linux, AIX, Solaris, and HP-UX physical environments as well as some of their virtual variants (for Linux, AIX, and Solaris).

Enterprise Strategy Group | **Getting to the bigger truth.**

20 Asylum Street | Milford, MA 01757 | Tel:508.482.0188 Fax: 508.482.0218 | www.enterprisestrategygroup.com