Extending the Virtualization Infrastructure

on Dell PowerEdge Servers Using VMware ESX Server 3 Advanced Features

Virtualization, which allows several operating systems to run simultaneously on a single physical server, is a key technology for achieving a scalable enterprise. Ninth-generation Dell™ PowerEdge™ servers also can help improve performance and scalability in enterprise data centers, especially when combined with VMware® ESX Server™ 3 virtualization software.

BY DAVID SCHMIDT, SCOTT STANFORD, AND BALASUBRAMANIAN CHANDRASEKARAN

Related Categories:

Dell ninth-generation servers

Dell PowerEdge servers

Virtual data center

Virtualization

VMware

Visit www.dell.com/powersolutions for the complete category index.

VMware ESX Server 3 introduces features and capabilities not available in previous versions of this virtualization platform, including expanded support for storage technologies, advanced guest OS features, and other performance enhancements. When combined with ninth-generation Dell PowerEdge servers—which offer enhanced processor, memory, and connectivity options—these virtualization capabilities can be further extended and can help build a highly scalable data center.

Features introduced in VMware ESX Server 3

The features introduced in VMware ESX Server 3—support for Internet SCSI (iSCSI) and network attached storage (NAS), VMware ESX Server File System version 3 (VMFS-3), the four-way VMware Virtual SMP™ feature, 16 GB memory for virtual machines (VMs), and experimental support for 64-bit guest operating systems—make a compelling case for implementing virtualization in data centers, branch offices, and remote field locations. These features allow large, resource-intensive applications to be hosted on virtual platforms, particularly applications that require Physical Address Extension[1] or hugemem[2] support for optimal performance. When combined and appropriately leveraged, these features can enable a robust, high-performance, and scalable virtualization framework at the individual ESX Server level.

Internet SCSI and network attached storage support

Past versions of ESX Server supported locally attached storage and Fibre Channel storage area networks (SANs)

[1] For more information about Physical Address Extension, see "Planning Considerations for Intel Extended Memory 64 Technology on Servers and Workstations" by John Coombs and John Fruehe in *Dell Power Solutions,* June 2004, www.dell.com/downloads/global/power/ps2q04-018.pdf.

[2] For more information about hugemem, see "RHEL4 x86 kernel with more than 4 GB of RAM" by Matt Domsch, e-mail to *Linux-PowerEdge* mailing list, August 20, 2005, lists.us.dell.com/pipermail/linux-poweredge/2005-August/022327.html.

for hosting and storing VM configuration files, guest OS virtual disk files, and application-level data within the guest OS flat file. To take advantage of the advanced VMware VMotion™ feature, previous-generation ESX Server hosts required shared storage supported by SANs.

In addition to supporting direct attach storage and Fibre Channel SANs, ESX Server 3 introduces support for iSCSI SANs and NAS devices. Both of these technologies provide storage across traditional Ethernet networks, which helps simplify deployment and management of storage fabrics and provides system administrators and storage architects with flexibility when designing, deploying, and upgrading ESX Server storage infrastructures. Although the initial list of supported NAS and iSCSI devices is limited, these technologies have the potential to deliver a new paradigm of enterprise storage.

Although still considered a new technology, iSCSI shows great promise as an enterprise-class storage fabric because of its cost-effectiveness and flexibility. Advanced Ethernet technologies and future iSCSI storage products will likely enhance the performance of iSCSI networks and lead to widespread adoption of this technology in the data center, where it can seamlessly scale and support the dynamics introduced by virtualization.

In addition to supporting direct attach storage and Fibre Channel SANs, ESX Server 3 introduces support for iSCSI SANs and NAS devices.

Like iSCSI, NAS provides flexible, cost-effective, high-volume storage. One NAS feature not commonly found in traditional monolithic file systems or storage devices is multi-protocol and multi–file system support. NAS devices provide simultaneous support for industry-standard file systems such as Network File System and Common Internet File System, making these multifunctional storage devices well suited for ESX Server, Linux®, UNIX®, or Microsoft® Windows® environments. In addition, many NAS systems can be seamlessly integrated into Fibre Channel SANs, serving as front-end gateways for ESX Server hosts that need to connect to back-end SAN disk resources over an IP-based transport layer.

VMware ESX Server File System version 3

VMFS-3 allows ESX Server 3 hosts to manage the entire VM on a single file system. In previous ESX Server versions, the VM swap file and configuration file were stored in a separate file system; VMFS-3 supports a directory structure in the file system, and all configuration and swap files are stored in the data store. This structure allows every component to be located on a single enterprise data store, which enables VMware VirtualCenter features such as Distributed Resource Scheduler and VMware High Availability to operate at the file system level.

VMware Virtual SMP

Four-way Virtual SMP is another key feature introduced with ESX Server 3. This feature provides support for four-processor symmetric multiprocessing (SMP) VMs, allowing guest operating systems to utilize up to four virtual processors. As in previous versions of ESX Server, Virtual SMP is enabled and scheduled at the VMkernel level for guest operating systems supporting applications that can benefit from or require SMP capabilities. Software that can benefit from four-way Virtual SMP include database, messaging, Web, streaming media, and other multi-threaded applications whose workloads are either processor-centric or most efficient when running in an SMP environment.

16 GB memory for virtual machines

In ESX Server 3, up to 16 GB of memory can be allocated to a VM, increasing the amount of memory that can be allocated to an individual VM by more than 75 percent compared with ESX Server 2.*x*. For applications such as databases that are memory sensitive, this increase can help yield substantial performance improvements, because much of the application-level workload can be contained within a VM's system memory. Because frequently requested application data can be held in a VM's memory, fewer calls are necessary to the VM virtual disk subsystem—meaning that in a VMware ESX Server 3 environment, the ESX Server VMkernel intercepts, translates, and manages fewer calls to the underlying physical disk controller and attached disk devices compared to an environment with less VM memory. For memory-sensitive applications, the end result is an efficient use of physical host-server system memory, reduced guest OS requests for data from the virtualized disk subsystem, and, most important, fast response times for database queries from end users.

Support for 64-bit guest operating systems

In the physical server environment, 64-bit processors and memory subsystems allow operating systems and applications to take advantage of wider and more efficient paths to system memory and processors compared to 32-bit hardware. As a result, more work can be done in the system memory than is possible on a 32-bit platform, resulting in fewer disk accesses and larger chunks of data processed per processor cycle.[3]

[3] For more information about the benefits of 64-bit processors and efficiencies gained by keeping work in system memory, see "Evaluating the Performance of 64-bit Oracle9*i* Database Release 2 on the Intel EM64T-based Platform" by Ramesh Radhakrishnan, Ph.D., in *Dell Power Solutions,* October 2004, www.dell.com/downloads/global/power/ps4q04-20040167-Radhakrishnan.pdf; "The Effect of L3 Cache Size on MMB2 Workloads" by Scott Stanford in *Dell Power Solutions,* February 2003, www1.us.dell.com/content/topics/global.aspx/power/en/ps1q03_stanford?c=us&cs=555&l=en&s=biz; and "VMware ESX Server Performance on Dell PowerEdge 2850 and PowerEdge 6850 Servers" by Todd Muirhead; Dave Jaffe, Ph.D.; and Scott Stanford in *Dell Power Solutions,* February 2006, www.dell.com/downloads/global/power/ps1q06-20050312-Muirhead.pdf.

The concept of wider data paths and more work per processor cycle carries over into the virtualized infrastructure with ESX Server 3. Guest operating systems and applications that are 64-bit aware can take advantage of the wider data paths on 64-bit platforms. For ESX Server 3, 64-bit guest OS support is experimental. Application, OS, and database development, testing, and quality assurance activities are no longer limited to the 32-bit OS or application realm.

A 32-bit OS employs methods that enable 64-bit-like access for some applications, but do so with a potential performance penalty given the overhead required to remap and present system memory beyond the 4 GB memory level to applications as if it were part of the normal 32-bit memory address space. When formal, nonexperimental support for 64-bit guest operating systems is offered in future ESX Server versions, production-level workloads should be able to run in a complete 64-bit stack.

Virtual data centers with VirtualCenter 2.0

Data centers are often viewed as collections of individual or isolated processing elements in which software services or tasks are tied to specific hardware to achieve guaranteed service levels. Virtualization breaks this tight bond between software and hardware by mobilizing the software components. This mobilization is accomplished primarily through the migration of VMs within the data center. Administrators can use ESX Server 3 features to build virtual data centers on pools of versatile, scalable ESX Server hosts.

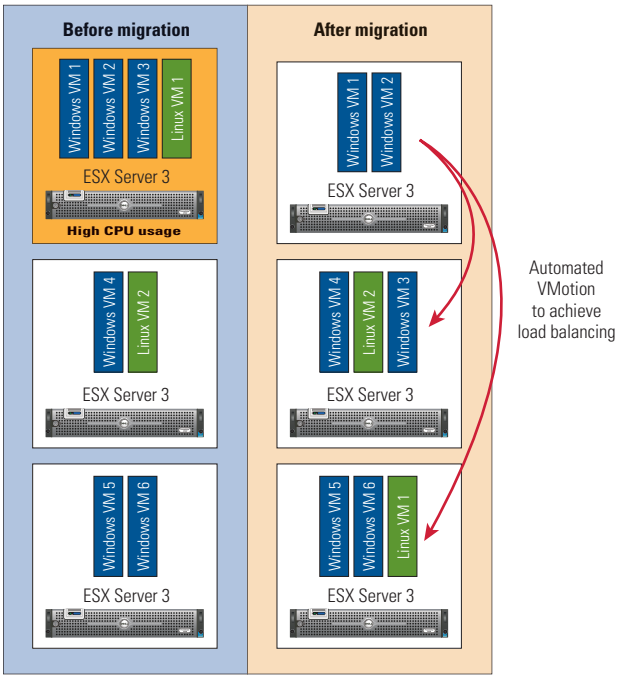

Figure 1. VMware Distributed Resource Scheduler and VMotion migrating virtual machines to a new location to achieve load balancing

In previous versions of VirtualCenter and ESX Server, VMotion events were primarily performed manually. To help improve on some of the manual processes that inherently limit dynamic systems reallocation in data centers, VirtualCenter 2.0 introduces a feature called Distributed Resource Scheduler (DRS). DRS provides a mechanism by which VMs can automatically migrate to a new location to achieve load balancing or avoid resource contention, as shown in Figure 1. DRS enables system administrators and architects to design application and workload containers on resource pools composed of VMs.

Because VMs are not tied to a particular server, DRS-based policies and rules enable VMs to autonomously and automatically migrate as data-center resource demands ebb and flow throughout the day or work cycle. Some key benefits provided by this autonomous and automatic migration feature are as follows:

- Physical data center resources can be used efficiently.
- Granular service-level agreement (SLA) guarantees can be achieved by configuring the VM resource pools for permanent or temporal assignment to specific hardware resource pools using an affinitization model or policy-based migration scheduling.
- Administrators can view the data center as a pool of resources instead of a collection of discrete servers.
- Where the software service is running becomes less relevant than it was without DRS.
- How to design and properly configure VM and hardware resource pools becomes more relevant than it was without DRS.
- Chief information officers have a robust and agile tool set with which to meet ever-changing business demands or strategic high-level IT initiatives.

Another key feature in VirtualCenter 2.0 is VMware High Availability (HA). VMware HA detects server failures and restarts the VMs on other servers, as shown in Figure 2. This method is similar to the uptime levels provided by Microsoft Cluster Server (MSCS) software, except that DRS can restart the VM on any other available ESX Server host in the appropriate hardware resource pool or VMware HA cluster. MSCS can be used for mission-critical applications, whereas DRS can be used for applications that can withstand minor downtime.

DRS in some configurations may be less complicated to set up than MSCS. Note that even though DRS is a VirtualCenter feature, once set up it can work independently of VirtualCenter. The ability to function independently is similar to an MSCS deployment model known as a majority node set server cluster.[4] For VMware HA, this

[4] For more information about majority node set server clusters, visit technet2.microsoft.com/WindowsServer/en/Library/c20dd042-5d52-49b2-889f-f163e0e112751033.mspx.

ability means that any failure in the VirtualCenter infrastructure does not affect the ability of servers to work together to achieve high availability.[5] Both DRS and VMware HA are available as add-on services to VirtualCenter 2.0, and when used effectively, can help IT organizations achieve high levels of efficiency and flexibility in data centers.

Features introduced in ninth-generation Dell PowerEdge servers

Supporting the resource pools, dynamic provisioning, and flexibility offered by VirtualCenter 2.0 DRS and VMware HA, and the robust ESX Server 3 features at the individual compute-node hardware layer, are ninth-generation Dell PowerEdge servers.

Dell PowerEdge 1950, PowerEdge 2950, and PowerEdge 6850 servers include such advanced features as virtualization-specific processor technology, large memory support, high-bandwidth peripheral interconnect buses, and high-density local storage. These advanced features, although appropriate for any supported OS, are particularly well suited to enabling stable, dense, high-performance ESX Server 3 hosts that can easily become part of large resource pools in the VirtualCenter 2.0 virtual infrastructure.

Processors with Intel Virtualization Technology

Ninth-generation PowerEdge servers offer both the Intel® Xeon® 5000 series and Intel Xeon 5100 series dual-core processors and associated chipsets. Intel Xeon processors in Dell ninth-generation servers support dual-core processing—multiple physical processing units per processor die—and enable the Intel Virtualization Technology (VT) hardware feature.

VT provides virtualization-specific instructions at the processor level that can be used by the VM monitor for context-switching operations between VMs. This avoids unnecessary binary translation of the privileged instructions executed by the VMs. VT is required for 64-bit guest operating systems.

Large memory support

Ninth-generation PowerEdge servers support larger memory capacity than eighth-generation PowerEdge systems. The PowerEdge 2950, for example, supports up to 32 GB of memory in an 8 × 4 GB fully buffered dual in-line memory module (DIMM)[6] configuration. The benefits of large system memory amounts in a 2U system such as the PowerEdge 2950 for ESX Server 3 deployments include the following:

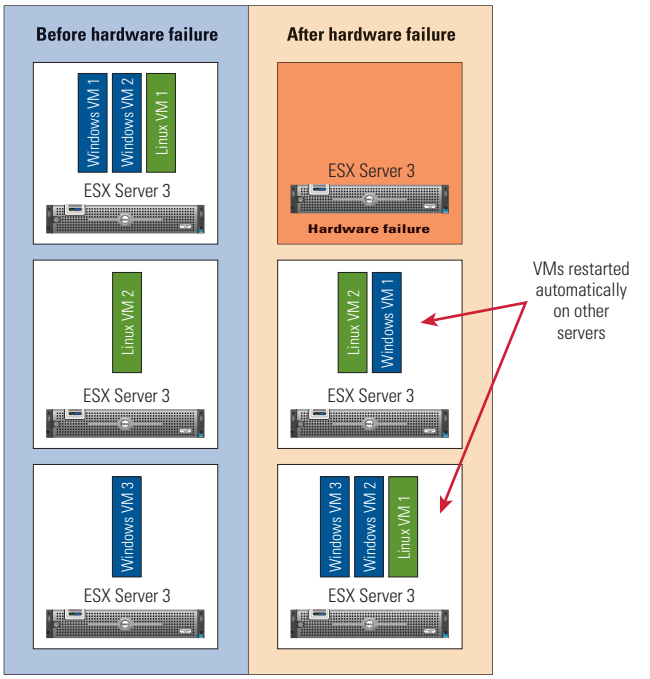

Figure 2. VMware High Availability Services restarting virtual machines on other servers following a hardware failure

- Support for more than 30 VMs, assuming a 1 GB active memory footprint per VM
- ESX Server host memory-system oversubscription models that take advantage of memory ballooning features provided by ESX Server and VMtools drivers

Ninth-generation PowerEdge systems with ESX Server 3 hosts also offer error-correction code (ECC) and single-device data-correction (SDDC)[7] features. ECC and SDDC implement single- and multi-bit memory-level error checking and correction functionality. These features are critical components designed to help VMs run in a reliable and stable memory subsystem.

PCI Express

Ninth-generation PowerEdge servers support port-dense, high-throughput, and low-latency PCI Express–based network interface card and host bus adapter configurations. High-performing PCI Express devices—with greater port density and higher capacity and throughput capabilities compared with previous-generation

[5] For more information about robust VirtualCenter infrastructure configurations, see "Architectural Considerations for Creating High-Availability VMware VirtualCenter Infrastructures" by Scott Stanford, Simone Shumate, and Balasubramanian Chandrasekaran in *Dell Power Solutions,* November 2005, www.dell.com/downloads/global/power/ps4q05-20050285-Stanford.pdf; and the VMware white paper "VMware VirtualCenter Technical Best Practices" at www.vmware.com/pdf/vc_technical_best.pdf.

[6] For more information about fully buffered DIMMs, see "Fully-Buffered DIMM Technology Moves Enterprise Platforms to the Next Level" by Jon Haas and Pete Vogt in *Technology@Intel Magazine,* March 2005, www.intel.com/technology/magazine/computing/Fully-buffered-DIMM-0305.htm.

[7] For more information about SDDC, visit www.intel.com/design/chipsets/applnots/29227401.pdf.

PCI technologies—allow ESX Server 3 to support larger numbers of virtual network switches and virtual SCSI devices and ultimately more VM-based I/O traffic than ESX Server hosts configured with previous-generation PCI technologies. Because more VMs can access the PCI Express–based devices through the virtualized hardware presented by ESX Server without incurring I/O bandwidth limitations found in legacy PCI-enabled buses, administrators can theoretically deploy more VMs per ESX Server host, assuming that application and SLA metrics can be met while hosting a larger number of VMs than was possible with eighth-generation servers.

Serial Attached SCSI

The Serial Attached SCSI (SAS) protocol is designed to support full-duplex transfer rates of up to 3 Gbps. SAS controller small-form factors are well-suited for use as on-board server boot controllers or as add-in PCI Express–based cards to support external storage devices. In addition, SAS controllers support the use of smaller disk-drive form factors than traditional parallel SCSI devices. Ninth-generation PowerEdge servers leverage these features and offer a variety of high-density disk drives that can be configured in fault-tolerant yet high-performance configurations. SAS storage configurations are particularly well suited for deploying ESX Server configurations that require high capacity or performing locally attached VMFS-3 or ESX Server boot partitions.

By combining ESX Server 3 with ninth-generation PowerEdge servers, administrators can use virtualization to enhance data center flexibility, reliability, and scalability.

Dell OpenManage 5

ESX Server 3 supports Dell OpenManage™ 5 systems management software. Dell OpenManage is a suite of systems management applications for managing Dell servers. Dell OpenManage 5 supports ninth-generation PowerEdge servers, with management support for SAS storage controllers and the next generation of the Dell Remote Access Controller.

As in previous versions of ESX Server, the ESX Server 3 service console supports Red Hat™ Package Manager (RPM™)–based installation of Dell OpenManage 5. After installing the software, administrators can access systems management information through the Dell OpenManage Server Administrator interface. Administrators can also use the Intelligent Platform Management Interface to control PowerEdge servers.

Dell certified configurations

Dell designs, tests, validates, and certifies Dell PowerEdge, Dell PowerVault™, and Dell/EMC storage configurations to maintain precise alignment with Dell's fundamental principles of customer focus, supportability, simplicity, and reliability. ESX Server software and VirtualCenter can be factory-installed on Dell systems. In addition, installation guides, systems management papers, white papers, and best practices guides for Dell-VMware configurations can be found at www.dell.com/vmware. The resources on this Web site are updated based on customer feedback and as new options become available for supported configurations.

Flexible, scalable virtual data centers

VMware ESX Server 3 offers advanced features designed to enable a dynamic and scalable enterprise data center, and the advanced industry-standard features of ninth-generation Dell PowerEdge servers can help improve data center performance and reliability. In addition, ninth-generation PowerEdge server hardware with Intel Virtualization Technology can align with ESX Server 3 features such as 16 GB memory support and experimental 64-bit guest OS support. By combining ESX Server 3 with ninth-generation PowerEdge servers, administrators can use virtualization to enhance data center flexibility, reliability, and scalability.

David Schmidt manages the Virtualization Solutions Engineering team within the Dell Global Solutions Engineering organization. Previously, David worked as a software developer in the systems management group, where he worked on the Dell OpenManage Deployment Toolkit. David has a B.S. in Computer Engineering from Texas A&M University.

Scott Stanford manages the Custom Solutions Engineering team within the Dell Global Solutions Engineering organization. Previously, he was a systems engineer in the Virtualization Solutions Engineering team. He has a B.S. from Texas A&M University and an M.S. in Community and Regional Planning from The University of Texas at Austin, and he is pursuing an M.S. in Computer Information Systems at St. Edward's University.

Balasubramanian Chandrasekaran is a systems engineer in the Virtualization Solutions Engineering team within the Dell Global Solutions Engineering organization. His research interests include virtualization of data centers, high-speed interconnects, and high-performance computing. Balasubramanian has an M.S. in Computer Science from The Ohio State University.

FOR MORE INFORMATION

Dell and VMware:
www.dell.com/vmware

