

Introducing VMware ESX Server, VirtualCenter, and VMotion on Dell PowerEdge Servers

VMware® ESX Server™ software enables administrators to provision multiple independent virtual machines on the same physical server. Dell engineers tested VMware ESX Server, VirtualCenter virtual machine management software, and VMotion™ virtual machine migration technology on Dell™ PowerEdge™ servers to illustrate how virtual machines can be moved from one physical server to another while processing heavy production loads.

BY DAVE JAFFE, PH.D.; TODD MUIRHEAD; AND FELIPE PAYET

IT managers today face a number of challenges as they are pushed to do more with less: improving service delivery, decreasing server sprawl, increasing system utilization, and making IT resources more flexible. VMware® server virtualization software running on Dell™ PowerEdge™ servers can help address these challenges.

VMware ESX Server™ software enables administrators to create multiple *virtual machines* on a single Intel® processor-based server, where each virtual machine can run a separate operating system (OS) and applications. VMware VirtualCenter provides centralized virtual machine monitoring and management from an easy-to-use graphical user interface (GUI). VMotion™ virtual machine migration technology enables administrators to move a running virtual machine from one physical server to another.

The Dell and VMware approach targets specific workload deployments in which server virtualization can offer the most value:

- **Test and development environments:** Virtualization can help administrators consolidate multiple

test and development servers onto fewer physical servers without sacrificing flexibility or functionality.

- **Application consolidation:** Virtualization enables administrators to consolidate applications from underutilized systems onto fewer physical servers, helping to simplify systems management and lower total cost of ownership (TCO) without compromising stability or security.

By deploying VMware ESX Server on multiple two- or four-processor servers and leveraging VMware VMotion, administrators may achieve several benefits not available on deployments that comprise a single server using eight or more processors:

- **Risk mitigation:** Virtual machines distributed among smaller servers can mitigate the impact of a hardware failure. In comparison, the failure of a single larger system would affect all virtual machines hosted by that one server.

Disk enclosures	Disks														
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
First DAE2											RAID-5: data staging				
Second DAE2	RAID-5: virtual machine (VM) boot drives														
Third DAE2	RAID-5: SQL data 1 (VM 1)					RAID-5: SQL data 2 (VM 1)					RAID-1: Logs (VM 1)		RAID-1: SnapView cache		Hot spare
Fourth DAE2	RAID-5: SQL data 1 (VM 2)					RAID-5: SQL data 2 (VM 2)					RAID-1: Logs (VM 2)				

Figure 1. Organization of LUNs on the Dell/EMC CX600 storage array

- **Expansion flexibility:** Deployments based on smaller, industry-standard building blocks permit a modular approach to expandability, whereby organizations can add incremental capacity using two- and four-processor servers instead of eight-processor and larger systems.
- **Operational flexibility:** A VMware deployment based on multiple Dell servers allows the live migration of virtual machines from one physical server to another using VMotion technology. This approach enables administrators to respond quickly to changes in workload demand and perform hardware upgrades or maintenance, all with minimal impact to workload delivery.

To evaluate Dell servers as a platform for server virtualization, in December 2003 a team of Dell engineers tested the performance of ESX Server software on the four-processor PowerEdge 6650 server using a Dell/EMC® storage area network (SAN). The Dell test team built an application that models an online DVD store on two instances of Microsoft® SQL Server™ 2000 Enterprise Edition. These database instances were deployed as virtual machines on two separate PowerEdge 6650 servers. One database instance received orders, and the other generated financial reports based on the order data. To determine whether virtual machines running heavy loads in a production environment could be moved without service interruption, the Dell team moved the virtual machine hosting the order entry database from one physical server to the other while the database was processing 100 orders per second—with no loss of transactions and only a slight rise in response time.

Setting up the hardware for the test environment

The two 4U Dell PowerEdge 6650 servers were each configured with VMware ESX Server 2.0.1 and four Intel® Xeon™ processors MP at 2.8 GHz with 2 MB of level 3 (L3) cache and 4 GB of RAM. Each PowerEdge 6650 server used a PowerEdge Expandable RAID Controller 3, Dual Channel (PERC 3/DC) and an Intel PRO/1000XT Gigabit Ethernet¹ network interface card (NIC) in addition to two on-board Gigabit Ethernet NICs. The

three NICs allowed dedicated bandwidth for the ESX Server service console, the virtual machines, and the VMotion workload management software.

The PowerEdge 6650 servers were attached to the Dell/EMC SAN by a QLogic® 2340 Fibre Channel host bus adapter (HBA). A Dell/EMC CX600 storage array was attached to the SAN to provide shared storage. The test team assigned 38 of the 150 drives attached to the CX600 for the VMware environment. The basic configuration of the CX600 storage array was as follows:

- **Disk enclosures:** Four Dell/EMC DAE2 disk array enclosures
- **Disks:** Thirty-eight 73 GB disks at 10,000 rpm
- **Logical storage units (LUNs):** One 6-disk RAID-5 LUN for the virtual machine boot drive, four 5-disk RAID-5 LUNs for database data storage, two 2-disk RAID-1 LUNs for database logs, one 5-disk RAID-5 LUN for temporary data staging before loading, one 2-disk RAID-1 split into two LUNs for an EMC SnapView™ storage management software cache, and one hot-spare disk
- **Software:** EMC Navisphere® Manager, EMC Access Logix™, and SnapView

One LUN on the Dell/EMC CX600 array was used to stage the data that was loaded into the database (see Figure 1). Using the snapshot capability of the SnapView software, the test team created a second copy of this data so that both virtual machines could load the data simultaneously.

Dell engineers used a PowerEdge 2650 server to produce a transaction load to run against the databases that were installed in the virtual machines on the two PowerEdge 6650 servers (see Figure 2). All servers, including the PowerEdge 6650 servers, were connected to a Dell PowerConnect™ 5224 Gigabit Ethernet switch for network connectivity. Using a Brocade® Fibre Channel switch, the test team also attached the PowerEdge 6650 servers to the Dell/EMC CX600 storage array.

All storage for the ESX Server-based virtual machines resided on the SAN, and each virtual machine was configured with its own

¹ This term does not connote an actual operating speed of 1 Gbps. For high-speed transmission, connection to a Gigabit Ethernet server and network infrastructure is required.

boot drive as well as two data drives and one log drive. When testers moved a virtual machine from one physical server to the other, only the RAM contents of the migrating virtual machine moved with it to the new physical hardware. Both servers already had access to storage, which was shared on the SAN.

Setting up the software for the test environment

The two VMware software products used for the Dell test were ESX Server and VirtualCenter. ESX Server has its own kernel that runs directly on the hardware and hosts virtual machines, enabling multiple virtual machines to run at the same time on the same hardware. VirtualCenter is a console application through which administrators can monitor and control ESX Server installations—and the virtual machines running on them—from a central location across multiple Dell servers.

Installing and configuring ESX Server

The test team configured the internal drives on the PowerEdge 6650 servers as RAID-1. The QLogic HBA was disconnected from the SAN during the initial stage of the ESX Server installation. To install ESX Server, Dell engineers booted from the ESX Server CD and answered the installation questions concerning partitioning of the local drives, the ESX Server host name, IP address, Domain Name System (DNS) server, gateway address, and initial root password. The team copied all necessary files from the installation CD and then rebooted the system.

To complete the installation of the ESX Server software—and for most administration and configuration tasks—the team accessed the ESX Server service console remotely through a Web browser. Following the initial installation stage, when administrators access the ESX Server for the first time through a Web browser, the software presents a series of configuration steps. These steps include installing the ESX Server license and configuring all hardware on the server that would be used by either the service console or virtual machines. The service console portion of each ESX Server installation requires a dedicated NIC. Dell recommends that the virtual machines also use one or more dedicated NICs per physical server. In this test, each virtual machine controlled its own HBA and all the SAN storage allocated to it.

After configuring the ESX Server hardware options, administrators must reboot the server. Just prior to rebooting, the Dell team connected the QLogic HBA into the SAN fabric and created a new zone on the switch for the newly connected server. Once the switch was correctly zoned, Dell engineers used Navisphere Manager—the management tool for the Dell/EMC CX600 storage array—to manually register the new host in the Connectivity Status screen. (Currently, no version of Navisphere Agent is available to register ESX Server automatically.) Once the registration was complete, the team used Navisphere Manager to create the necessary RAID groups, LUNs,

and storage groups. Administrators must assign all ESX Server systems expected to participate in VMotion virtual machine migrations to the same storage group.

Adding ESX Server servers to the VirtualCenter service console

VirtualCenter, a Microsoft Windows®-based program, was installed on a separate PowerEdge 1750 system that served as the management node for the test configuration. Dell engineers added all ESX Server virtual machines to be managed by VirtualCenter to the VirtualCenter console using a simple Connect Host wizard, which prompted for the host name, user ID, and password of each system running ESX Server. Adding all the ESX Server systems to VirtualCenter enables administrators to perform management tasks—including cloning, template production, and VMotion virtual machine migration—from the VirtualCenter console for any virtual machines that reside on those ESX Server systems.

Creating the virtual machines

The Dell team used VirtualCenter to create a new virtual machine on the SAN, specifying the Microsoft Windows Server™ 2003 Enterprise Edition OS, a 10 GB hard disk, 1 GB of RAM, and two CPUs. (The symmetric multiprocessing, or SMP, feature of ESX Server allowed the virtual machine to use two physical CPUs.) VirtualCenter created a virtual machine ready for installation of the OS. The Dell team then booted the virtual machine from the ISO image of the Windows Server 2003 Enterprise Edition installation CD and installed the OS on the virtual machine. The database application was installed afterward. Dell engineers then created two clones of this virtual machine

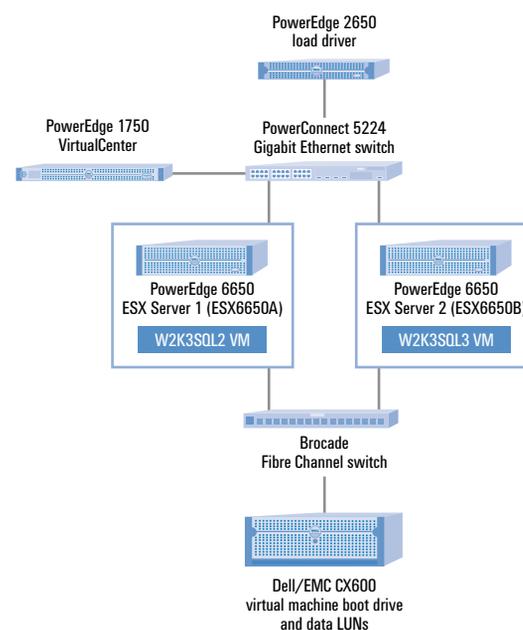


Figure 2. Configuration of servers and storage used for testing

master for use in testing. After the virtual machines were created, each was assigned additional hard disks for the data and logs of the database that resided on the CX600 storage array (see “Setting up the hardware for the test environment”).

Examining the test database application: An online DVD store

To demonstrate the advantages of running a large application on VMware ESX Server, Dell engineers created a 100 GB online DVD store, which they implemented as two replicated database instances, each running on its own virtual machine. One of the database instances handled the entry of new orders and replicated changes on a scheduled basis to the second database instance, which was used for generating financial reports.

The DVD store database consisted of a set of data tables organized according to a certain schema, as well as a set of stored procedures that did the actual work of managing the data in the database as orders were entered and reports requested. The database back end was designed to be driven from a Web-based middle tier, but because the focus of the Dell test was on the database servers, the back-end stored procedures were driven directly by custom programs written in the C programming language to simulate a Web-based middle tier.

Understanding the database schema

The DVD store comprised four main tables and one additional table (see Figure 3). The Customers table was prepopulated with 200 million customers: 100 million U.S.-based customers and 100 million customers from the rest of the world. The Orders table was prepopulated with 10 million orders per month, starting in January 2003 and ending in September 2003. The Orderlines table was prepopulated with an average of five items per order. The Products table contained 1 million DVD titles. An additional Categories table listed the 16 DVD categories. For the full DVD store database build script used in this test, visit *Dell Power Solutions* online at http://www.dell.com/magazines_extras.

Managing the database using stored procedures

The Dell team managed the DVD store database using seven stored procedures. The first two procedures were used during the login phase. For returning customers, the Login procedure retrieved the customer’s information—in particular, the CUSTOMERID. For new customers, the New_customer procedure created a new row in the Customers table containing the customer’s data.

Following the login phase, the customer might search for a DVD by category, actor, or title. These database functions were implemented by the Browse_by_category, Browse_by_actor, and Browse_by_title procedures, respectively. Finally, after the customer completed the selections, the Purchase procedure was called to complete the transaction. Additionally, the Rollup_by_category procedure calculated total sales by DVD category for the previous month,

Table	Columns	Number of rows
Customers	CUSTOMERID, FIRSTNAME, LASTNAME, ADDRESS1, ADDRESS2, CITY, STATE, ZIP, COUNTRY, REGION, EMAIL, PHONE, CREDITCARD, CREDITCARDEXPIRATION, USERNAME, PASSWORD, AGE, INCOME, GENDER	200 million
Orders	ORDERID, ORDERDATE, CUSTOMERID, NETAMOUNT, TAX, TOTALAMOUNT	90 million
Orderlines	ORDERLINEID, ORDERID, PROD_ID, QUANTITY, ORDERDATE	450 million
Products	PROD_ID, CATEGORY, TITLE, ACTOR, PRICE, QUAN_IN_STOCK, SPECIAL	1 million
Categories	CATEGORY, CATEGORYNAME	16

Figure 3. Database schema for online DVD store

quarter, and half-year periods. For the stored procedures, visit *Dell Power Solutions* online at http://www.dell.com/magazines_extras.

Using driver programs to model workloads

The Dell team wrote separate multithreaded driver programs to model the order entry, or *online transaction processing* (OLTP), workload as well as the report request workload. Each thread of the OLTP driver application connected to the database and made a series of stored procedure calls that simulated customers logging in, browsing, and purchasing. Because Dell engineers did not simulate customer think time or key time, the database connections were kept full—simulating a multitiered application in which a few connections are pooled and shared among Web servers that may be handling thousands of simultaneous customers. In this way, the test team achieved a realistic simulation of database activity without having to model thousands of customers.

Each thread of the OLTP driver program modeled a series of customers going through the entire sequence of logging in, browsing the catalog several ways, and purchasing selected items. Each completed customer sequence counted as a single order. The OLTP driver program measured order rates and the average response time to complete each order. Several tunable parameters were used to control the application (see Figure 4).

The report request driver program was similar to the OLTP driver program in that each thread connected to the database and started making stored procedure calls. Each thread made repeated calls to the Rollup_by_category stored procedure until reports for all 16 DVD categories were completed. In each test, eight simultaneous reports were run.

Moving a virtual machine under heavy load

To demonstrate the capability of VMware software to move virtual servers around a farm of physical servers, the Dell team used the VMware VMotion add-on to VirtualCenter, which enables administrators to move a virtual machine from one physical server running

ESX Server to another. The migration was performed while the virtual machine was running the DVD store database under a heavy stress load of 100 orders per second. In a live production environment, such a move might be required to balance workloads among computing resources, perform routine maintenance on a server, or respond to an alert that a server parameter such as temperature had exceeded a warning threshold. In Figure 5, the VirtualCenter console shows the virtual machines in the test server farm.

At the start of the test, one node of the database replication group, W2K3SQL3, on physical PowerEdge 6650 server ESX6650B was handling approximately 100 orders per second with an average response time of 0.1 second. For the test, response time was defined as the total response time experienced by the simulated customer for the complete order transaction, including login time, browse time, and response time after the customer pressed the Submit button to purchase the order.

Dell engineers then started the second database system, W2K3SQL2, running on physical server ESX6650A, which began calculating sales by DVD category for eight separate categories. In addition, the test team set up the servers to replicate new orders from the W2K3SQL3 node to the W2K3SQL2 node once per day. The two virtual machines running database instances are shown in the ESX Server service console in Figure 6.

The Dell team started the order entry and the report request workloads against the two database instances, each instance running in a virtual machine on its own PowerEdge 6650 server. Each server achieved full speed—100 orders per minute, or eight simultaneous reports—using about 80 percent of the two CPUs dedicated to each virtual machine. The Dell team used the VMotion

Parameter	Description	Value(s) used in test
n_threads	Number of simultaneous connections to the database	10
warmup_time	Warm-up time before statistics are kept	1 minute
run_time	Runtime during which statistics are kept	Varied
pct_returning	Percent of customers who are returning	95 percent
pct_new	Percent of customers who are new	5 percent
n_browse_category	Number of searches based on category	Range: 1–3 Average: 2
n_browse_actor	Number of searches based on actor	Range: 1–3 Average: 2
n_browse_title	Number of searches based on title	Range: 1–3 Average: 2
n_line_items	Number of items purchased	Range: 1–9 Average: 5
net_amount	Total amount of purchase	Range: \$0.01–\$400.00 Average: \$200.00

Figure 4. OLTP driver parameters

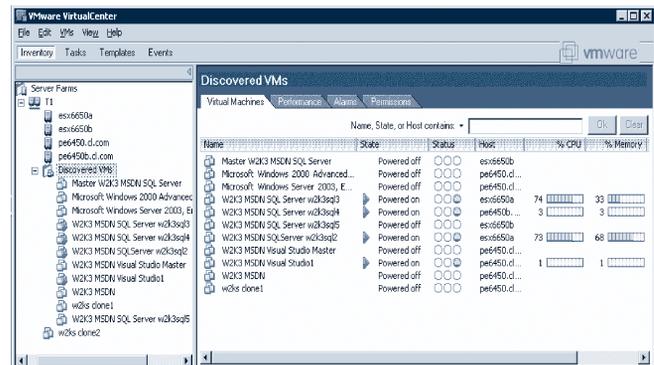


Figure 5. VMware VirtualCenter displaying the virtual machines in the test server farm

feature of VirtualCenter to move the virtual machine performing order entry (W2K3SQL3) from physical server ESX6650B to physical server ESX6650A, without stopping either the incoming orders or the sales calculations on W2K3SQL2. Figures 7 and 8 show the results of this migration.

As shown in Figure 8, for the first 25 seconds after the VMotion migration was initiated at 15:36:20, there was little impact on either throughput (orders per second, indicated in the top half of Figure 8) or response time (indicated in the bottom half of Figure 8) while VirtualCenter prepared for the move by initializing a new virtual machine on the target ESX Server and synchronizing the memory between the two. At about 15:36:45, the effects of the memory synchronization could be seen in the dropping throughput and increasing response time.

The actual move occurred at 15:37:08, and the response time reached a maximum of 2.572 seconds while the order handling paused for approximately two seconds. Immediately after the move, the throughput and response time rapidly returned to close to their previous levels. The target ESX Server CPU

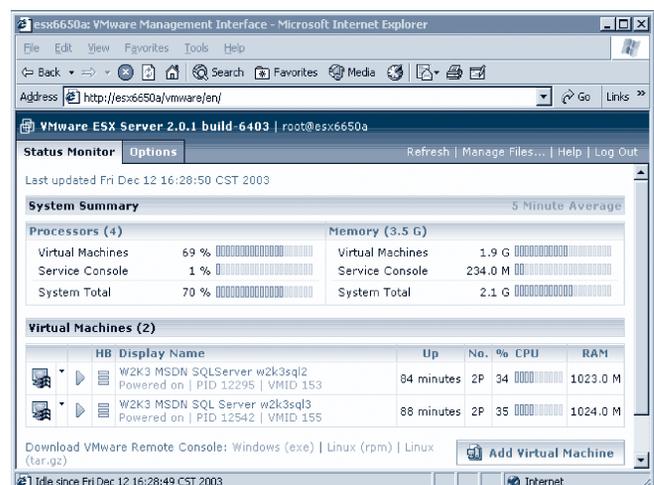


Figure 6. ESX Server service console showing the two virtual machines running database instances

Time	New orders completed per second	Average response time (seconds)	Maximum response time (seconds)
Before VMotion migration	103	0.098	0.201
During VMotion migration	80	0.139	2.572
After VMotion migration	93	0.109	0.492

Figure 7. Performance results before, during, and after VMotion migration of virtual machine running database application under heavy load

utilization rose to about 80 percent as both virtual machines ran on the target server, using two CPUs each. The throughput decreased slightly from the pre-VMotion level but was still high enough to handle 300,000 orders per hour while the first system was being repaired or upgraded.

Using virtual machine migrations to increase operational flexibility

The test findings in this article indicate that ESX Server software running on Dell PowerEdge servers with Dell/EMC SAN storage can provide a robust platform for server virtualization. In the Dell test discussed in this article, two new virtual machines were rapidly cloned from a single master and then used to implement a large online DVD store with one server handling new orders and then replicating the orders to the second server for reporting.

Using VMware VMotion workload management software, an add-on to VMware VirtualCenter, testers demonstrated that a virtual machine handling 100 orders per second could be moved from one physical server to another in less than a minute without stopping the database application and without losing any transactions. Test findings indicate that the slight increase in response time would be nearly imperceptible to the end user. Although the virtual machine migration took 48 seconds, the increased response time of less than three seconds from the end user’s perspective—at the point when the virtual machine actually switched from one physical server to the other—was experienced for only a second or two.

Deploying virtual machines on Intel processor-based servers can help IT organizations scale out cost-effectively and respond quickly and flexibly to changes in workload demand. The virtual server approach to IT management also can provide a convenient way to upgrade and maintain production servers in real time, without interrupting service to business-critical applications. In addition, VMware virtual machines running on industry-standard Dell servers can improve system availability and fault tolerance by avoiding a single point of hardware failure, as opposed to a single larger server. ☞

Acknowledgments

The authors would like to thank Craig Lowery, Tim Abels, and Wenlong Xu of the Scalable Enterprise Computing team at Dell for valuable discussions.

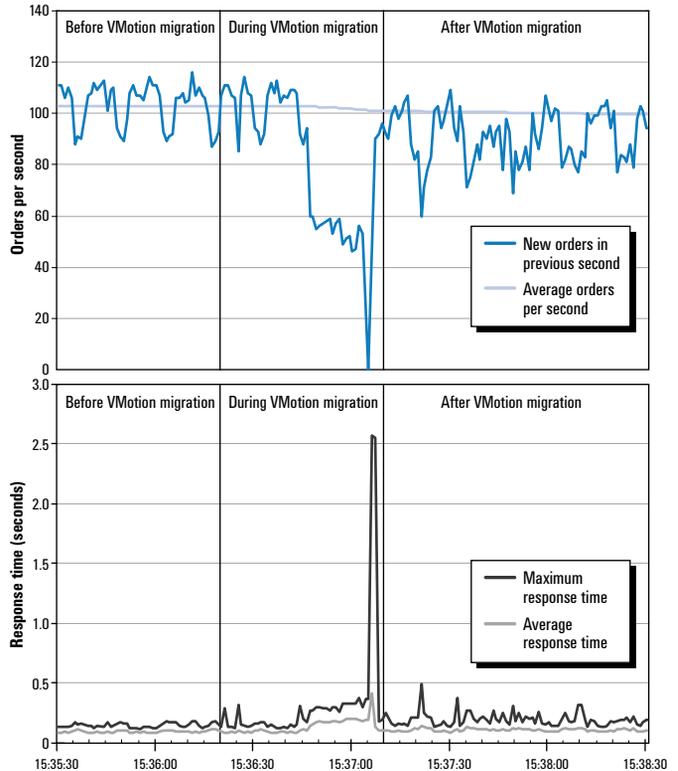


Figure 8. Throughput and response times before, during, and after VMotion migration

Dave Jaffe, Ph.D. (dave_jaffe@dell.com) is a senior consultant on the Dell Technology Showcase team who specializes in cross-platform solutions. Previously, Dave worked in the Dell Server Performance Lab, where he led the team responsible for Transaction Processing Council (TPC) benchmarks. Before working at Dell, Dave spent 14 years at IBM in semiconductor processing, modeling, and testing, and in server and workstation performance. Dave has a Ph.D. in Chemistry from the University of California, San Diego, and a B.S. in Chemistry from Yale University.

Todd Muirhead (todd_muirhead@dell.com) is an engineering consultant on the Dell Technology Showcase team. He specializes in SANs and database systems. Todd has a B.A. in Computer Science from the University of North Texas and is Microsoft Certified Systems Engineer + Internet (MCSE+I) certified.

Felipe Payet (felipe_payet@dell.com) manages the Dell and VMware relationship within the Software Alliance Team of the Dell Enterprise Server Group. Previously, he worked in various product management, business development, and emerging technology marketing roles at Dell, Intel, and several start-ups. Felipe has a B.A. in Economics from Yale University and an M.B.A. from the Sloan School of Management at M.I.T.

FOR MORE INFORMATION

Dell and VMware:
<http://www.dell.com/vmware>