POWER SOLUTIONS

DELL.COM/PowerSolutions

NOVEMBER 2007 • $4.95

REDUCING STORAGE COMPLEXITY

INSIDE THIS ISSUE:

REDUCING STORAGE COMPLEXITY

Learn more: intel.com/go/10GbE

Intel® 10 Gigabit XF SR Server Adapter
Now available for Dell PowerEdge* Servers

NOVEMBER 2007

POWER SOLUTIONS

DELL.COM/PowerSolutions

COVER STORY

REDUCING STORAGE COMPLEXITY

By Eric Endebrock; Dave Jaffe, Ph.D.; and Kathryn White

Digital data is outstripping storage capacity and complex systems are failing to adapt under the strain as storage demands grow faster than IT budgets. The good news is that easy-to-use IP-based networked storage with built-in data protection and management capabilities promises to bridge the budget gap—and effectively complement traditional Fibre Channel–based storage area networks.

STORAGE CONSOLIDATION WITH THE DELL POWERVAULT MD3000i iSCSI ARRAY

By Dave Jaffe, Ph.D., and Kendra Matthews

The Dell PowerVault MD3000i Internet SCSI (iSCSI) array enables enterprises to easily consolidate storage for multiple servers. Dell engineers performed tests demonstrating that the PowerVault MD3000i provides sufficient throughput in a consolidated environment.

SIMPLIFY YOUR IT · page 54

THE DELL PATH TO INNOVATION AND GROWTH

By Joe Pollock

Enterprises of all sizes struggle with the ever-increasing complexity of IT. By helping simplify and streamline the requirements for ongoing operations and systems maintenance, the Dell IT simplification approach enables organizations to reclaim the time, money, and personnel needed to turn innovation into a daily business practice instead of a long-term business goal. For more information, visit **DELL.COM/Simplify**.

NOW PLAYING: DELL DATACENTER OF THE FUTURE PODCAST SERIES

Flexible computing may conjure up a confusing array of terms: thin clients, virtual clients, and bladed desktops, for example. However, it is not one technology but a holistic strategy—controlling data while providing flexible data access. "Flexible Computing: What Is It and How Will It Simplify IT?" explains flexible computing initiatives from Dell. For more information, visit **DELL.COM/Podcast**.

EDITOR'S COMMENTS

FEATURE SECTION: NEW HORIZONS IN ENTERPRISE STORAGE

SEE IT HERE FIRST

MORE ONLINE

DELL.COM/PowerSolutions

NEW HORIZONS IN ENTERPRISE STORAGE

Oracle Data Guard Deployment Using Dell/EMC CX3 Storage Replication

By Wendy Chen, Ujjwal Rajbhandari, and Kai Yu

Using Dell/EMC CX3 storage replication software when deploying Oracle Data Guard can help significantly reduce primary database downtime during the deployment process.

Parallel Rendering Technologies for HPC Clusters

By Li Ou, Ph.D.; Yung-Chin Fang; Onur Celebioglu; and Victor Mashayekhi, Ph.D.

Using parallel rendering technologies with clusters of high-performance computing (HPC) workstations configured with high-end graphics processors helps scale out graphics capabilities.

Utilizing Open MPI Middleware for HPC Clusters

By Toby Sebastian, Sanjay Lalwani, and Munira Hussain

Open MPI helps simplify the porting and execution of Message Passing Interface (MPI)–based applications and supports many HPC platforms, interconnects, and environments.

Migrating Remote Installation Services to Windows Deployment Services in Microsoft Windows Server 2003

By Sankari N., Manjunath Narayanan, and Gobinath K.

Microsoft Windows Deployment Services can be used for deploying the Windows Vista OS as well as the forthcoming Windows Server 2008 OS and a variety of legacy operating systems.

Using IPMItool Raw Commands for Remote Management of Dell PowerEdge Servers

By Seshadri N. and Raghavendra Babu

To help maximize the capabilities of the baseboard management controllers in Dell PowerEdge servers, administrators can access hardware functions using IPMItool raw commands.

Xen Virtualization in Red Hat Enterprise Linux 5

By Jose De la Rosa and Puneet Dhawan

The open source Xen virtualization hypervisor included in Red Hat Enterprise Linux 5 provides key components to help build dynamic, scalable virtualized environments.

Maintaining Dell Platforms with Dell Technical Updates

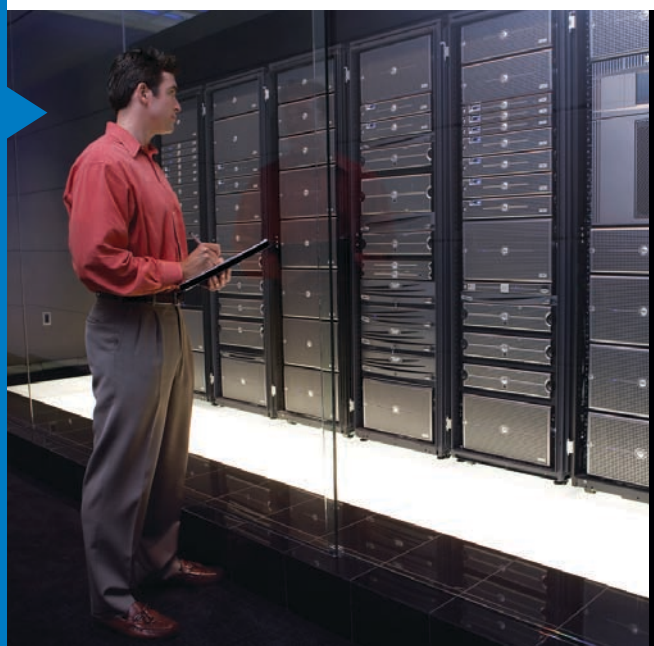
Firmware updates for Dell PowerEdge components and Dell PowerVault storage can help improve system functionality and minimize potential problems.

YOUR INFORMATION, YOUR WAY

The *Dell Power Solutions* articles on this page are available exclusively online at **DELL.COM/PowerSolutions**. Plus: Check the *Dell Power Solutions* Web site for our early-release articles, how-to's, implementation studies, and expert tips you won't find anywhere else.

Need specific subject matter at your fingertips? Search our **Related Categories** index and then visit full-text articles from current and back issues online. Can't wait? For real-time access to content as soon as it's posted, subscribe your favorite RSS (really simple syndication) reader to the feed at the *Dell Power Solutions* Web site or **DELL.COM/RSS**.

DEPARTMENTS



TALK BACK

We welcome your questions, comments, and suggestions. Please send your feedback to the *Dell Power Solutions* editorial team at **us_power_solutions@DELL.COM**.

ADVERTISER INDEX

DELL POWER SOLUTIONS

DELL.COM/PowerSolutions

EDITORIAL STAFF

Editor-in-Chief and Publisher • Tom Kolnowski
Managing Editor • Deb McDonald
Features Editor • Kathryn White
Associate Managing Editor • Liza Graffeo
Senior Editor • Jim Duncan
Editorial Assistant • Amy J. Parker
Art Director and Cover Designer • David Chan
Designer and Illustrator • Cynthia Webb

Contributing Writers • Mahmoud Ahmadian; Mike Apigian; Raghavendra Babu; Michael Bechauf; Max A. Benham; Richard Brown; Onur Celebioglu; Wendy Chen; David Christensen; Jose De la Rosa; Puneet Dhawan; Lauren Duda; Eric Endebrock; Yung-Chin Fang; Franklin Flint; Art Gilliland; John Gudmundson; Aziz Gulbeden; Scott Hanson; David Havener; Munira Hussain; Dave Jaffe, Ph.D.; Clay Johnson; Gobinath K.; Chethan Kumar; Sanjay Lalwani; Jacob Liberman; Robert Lusinsky; Victor Mashayekhi, Ph.D.; Kendra Matthews; Jaleel Mohammed; Todd Muirhead; William Muscato; Sankari N.; Seshadri N.; Bala Nagesh; Manjunath Narayanan; Shivaraj Nidoni; Li Ou, Ph.D.; Alvin Peng; David Pennington; Joe Pollock; Arun Rajan; Ujjwal Rajbhandari; Paul Rubin; Amina Saify; Toby Sebastian; Suman Kumar Singh; Sumner Stewart; Peter Swartz; Sreeram Vedantham; Travis Vigil; Kathryn White; Kong Yang; Patty Young; and Kai Yu

Contributing Photographers • Tony Bolding, Bryan Kuntz, Adran Matte, and Joey Pena

Online Design • Joi Chevalier
Special Inserts Manager • Kristin Kreisel

SUBSCRIPTIONS

Subscriptions are complimentary to qualified readers who complete the online subscription form. To sign up as a new subscriber, renew an existing subscription, change your address, or cancel your subscription, access the online Subscription Center forms at **DELL.COM/PowerSolutions**. For other subscription services, please e-mail **us_power_solutions@DELL.COM**.

ABOUT DELL

Dell Inc., headquartered in Round Rock, Texas, near Austin, listens to its customers and delivers innovative technology and services they trust and value. Uniquely enabled by its direct business model, Dell is a leading global systems and services company and No. 34 on the Fortune 500 list. For more information, visit our Web site at **DELL.COM**.

Dell Power Solutions is published in February, May, August, and November by Dell Inc., *Dell Power Solutions*, One Dell Way, Mail Stop RR5-03, Round Rock, TX 78682, U.S.A. No part of this publication may be reprinted or otherwise reproduced without permission from the editor-in-chief. Dell does not provide any warranty as to the accuracy of any information provided through *Dell Power Solutions*. Opinions expressed in this magazine may not be those of Dell. The information in this publication is subject to change without notice. Any reliance by the end user on the information contained herein is at the end user's risk. Dell will not be liable for information in any way, including but not limited to its accuracy or completeness. Dell does not accept responsibility for the advertising content of the magazine or for any claims, actions, or losses arising therefrom. Goods, services, and/or advertisements within this publication other than those of Dell are not endorsed by or in any way connected with Dell Inc.

Printed on recycled paper containing 10 percent post-consumer waste.

November 2007

STOCK NUMBER: DPS 2007-11 (NOVEMBER 07)

IN SEARCH OF IP STORAGE

Tom Kolnowski
Editor-in-Chief and Publisher
tom_kolnowski@DELL.COM
DELL.COM/PowerSolutions
DELL.COM/RSS

Internet Protocol (IP)—the basic network building block of the Internet—is apparently alive and quite well. A recent search of the Internet Engineering Task Force (IETF) Web site (www.ietf.org) revealed a steady stream of recent RFCs (Requests for Comments) around IP—the latest being that of RFC 5072: IP Version 6 over PPP, published in September 2007. RFCs are typically conceived and written by IETF working group members, in what amounts to a highly collaborative but much more tightly controlled process than that of a wiki.

The IETF remains the preeminent repository for all things IP, so it should come as no surprise that the first seeds of the Internet SCSI (iSCSI) protocol were sown here. RFC 3720: Internet Small Computer Systems Interface (iSCSI) was first published in April 2004. Since then, IP-based storage has emerged from the engineering labs and grown in a huge way, with iSCSI riding the wave of Gigabit Ethernet into a wider portion of the portfolios in many enterprise storage farms.

Back in the *Dell Power Solutions* realm, we have partnered with our Dell storage engineering teams to bring you a groundbreaking collection of articles on the subject of iSCSI. Entitled "iSCSI: Changing the Economics of Storage," the series launched in the May 2007 issue of *Dell Power Solutions*, with follow-on installments in August 2007 as well as this November 2007 issue. In the first installment, "Understanding iSCSI in Enterprise Environments" (DELL.COM/downloads/global/power/ps2q07-20070335-Vigil.pdf), we explored the fundamentals of iSCSI storage architecture and looked at various performance, manageability, and security aspects. The next installment, "Deploying iSCSI in Virtualized Data Centers" (DELL.COM/downloads/global/power/ps3q07-20070401-Baker.pdf), examined the expanding role of iSCSI storage in virtualized environments with a focus on deployment, management, and data protection aspects. And in the third installment, "Using iSCSI in Small and Medium Businesses" (DELL.COM/downloads/global/power/ps4q07-20070402-Vigil.pdf), we delve into how small and medium businesses, remote offices, and enterprise departments and workgroups can maximize their return on an iSCSI infrastructure investment. We hope you find relevant information in this primer as you plan the integration of iSCSI in your own data centers.

Starting on page 8, our cover story, "Reducing Storage Complexity," leads a comprehensive lineup of 11 articles on the evolving storage landscape—and explains how smart technology choices can help bridge the IT budget gap. This in-depth feature coverage provides a comprehensive guide for architecting your storage plans, including an evaluation of new IP-based storage options designed to complement existing Fibre Channel storage area networks. Visit DELL.COM/PowerSolutions to read it all online at our newly redesigned Web presence, where we have enhanced our RSS feeds (check out the My *Dell Power Solutions* page) and made it easier to access archived issues of the magazine.

"DELL AND EMC...BEST REPUTATION FOR PRODUCT QUALITY"

Nexus Energy Software (April 2007)
Brian Lenane, VP Client Services

Dell/EMC brings you the performance and reliability you need in a storage solution. Our systems are easy to deploy, manage, and upgrade so your IT team can spend less time on maintenance and more time on big ideas.

SIMPLIFY IT BY CALLING (877) 671-3355 OR VISITING DELL.COM/EMC

DELL | EMC²

REDUCING STORAGE COMPLEXITY

DIGITAL DATA IS OUTSTRIPPING STORAGE CAPACITY AND COMPLEX SYSTEMS ARE FAILING TO ADAPT UNDER THE STRAIN AS STORAGE DEMANDS GROW FASTER THAN IT BUDGETS. THE GOOD NEWS IS THAT EASY-TO-USE IP-BASED NETWORKED STORAGE WITH BUILT-IN DATA PROTECTION AND MANAGEMENT CAPABILITIES PROMISES TO BRIDGE THE BUDGET GAP—AND EFFECTIVELY COMPLEMENT TRADITIONAL FIBRE CHANNEL–BASED STORAGE AREA NETWORKS.

BY ERIC ENDEBROCK
DAVE JAFFE, PH.D.
KATHRYN WHITE

Related Categories:

Dell PowerVault storage
Dell/EMC storage
Fibre Channel
Internet SCSI (iSCSI)
Serial Attached SCSI (SAS)
Storage
Storage architecture
Storage area network (SAN)
Storage consolidation
Virtualization

Visit DELL.COM/PowerSolutions for the complete category index.

While data growth has exploded at the rate of terabytes, budgets have not. As high-resolution data capture increases relentlessly and enterprise applications spawn ever more sophisticated transactions, cost-efficient storage and unified management for diverse data types have become pressing concerns. In an attempt to keep up with government regulations and privacy standards, many organizations have hurriedly pieced together data management and security tools with expanded storage platforms.

However, the additional burden of complex, heterogeneous storage systems that have evolved through acquisition and merger or in the absence of a centralized IT strategy only complicates matters further. The ramifications of this burgeoning complexity are reaching far beyond the data center to threaten business growth and innovation.

Many executives are already feeling the pinch as storage requirements continue to grow faster than typical IT budgets (see Figure 1). To help avoid shortfalls in funding that might jeopardize support for ongoing service-level agreements, IT organizations must find ways to reduce storage complexity and plug the drain on valuable time and resources. By increasing operational efficiency, manageability, and flexibility, enterprises can help bridge the budget gap and free IT resources to focus on strategic business initiatives.[1]

[1] For more information about the Dell approach to reducing IT complexity, see "Simplify IT: The Dell Path to Innovation and Growth," by Joe Pollock, in *Dell Power Solutions,* November 2007, DELL.COM/downloads/global/power/ps4q07-20080130-SimplifyIT.pdf.

"Implementing a comprehensive storage strategy based on simple, cost-effective, and capable plug-and-play technologies can help reduce costs, increase efficiency, and ensure business continuity and compliance."

An effective storage simplification strategy begins at the moment data is created and enables seamless data protection and management through all stages of the information life cycle, from backup and restore to archiving and ultimately deletion. In the past, such a comprehensive scope was difficult for many organizations to achieve because sophisticated storage management tools and capabilities were almost exclusively the realm of high-end UNIX® and mainframe systems—and cost-prohibitive for many small and medium businesses (SMBs), remote offices, and enterprise departments and workgroups. Today, Internet SCSI (iSCSI)–based storage area network (SAN) technology, plug-and-play Ethernet infrastructure components, and integrated, out-of-the-box data protection and management capabilities put storage consolidation and networked storage environments within reach of enterprises of all sizes.

IP-BASED NETWORKED STORAGE CONVERGENCE CHANGES ECONOMIES OF SCALE

At the forefront of this major advance is iSCSI technology, which routes data packets through standard Ethernet networks. This advanced protocol is enabling iSCSI-based SANs to extend the benefits of storage consolidation and a shared storage environment to organizations challenged by the same IT pressures as large enterprises, but with smaller budgets and fewer IT administrators.[2]

Taking advantage of broad-based support for the open Ethernet protocol and many common operating systems, applications, and platforms—including the Microsoft® Windows® and Linux® operating systems as well as Microsoft Exchange, Microsoft SQL Server™, Oracle® Database 11*g*, and VMware® virtualization software[3]—iSCSI technology enables a significant reduction in the cost of entry for a networked storage environment compared with Fibre Channel–based SANs. Moreover, iSCSI helps simplify storage in virtualized data center environments by allowing storage resources to be mapped directly to virtual machines (VMs). For more information, see the "How iSCSI storage unlocks the power of virtual servers" sidebar in this article.

Because Ethernet is ubiquitous in today's IT infrastructure, iSCSI enables organizations to capitalize on existing expertise to deploy networked storage without the additional cost or special equipment and training that Fibre Channel can require. Solutions such as the Dell™ PowerVault™ MD3000i modular disk storage array are designed to perform as a full-fledged enterprise SAN. For more information, see the "iSCSI–based SANs in action" sidebar in this article.

In addition, converging networked storage technologies allow iSCSI to be incorporated into existing SAN and network attached storage (NAS) environments. For example, Dell/EMC CX3 series arrays support both iSCSI and Fibre Channel, and the Dell PowerVault NX1950 unified NAS system includes iSCSI functionality.[4] And, as 10 Gigabit Ethernet technology becomes available, iSCSI technology is expected to enable unification of the data center storage fabric.

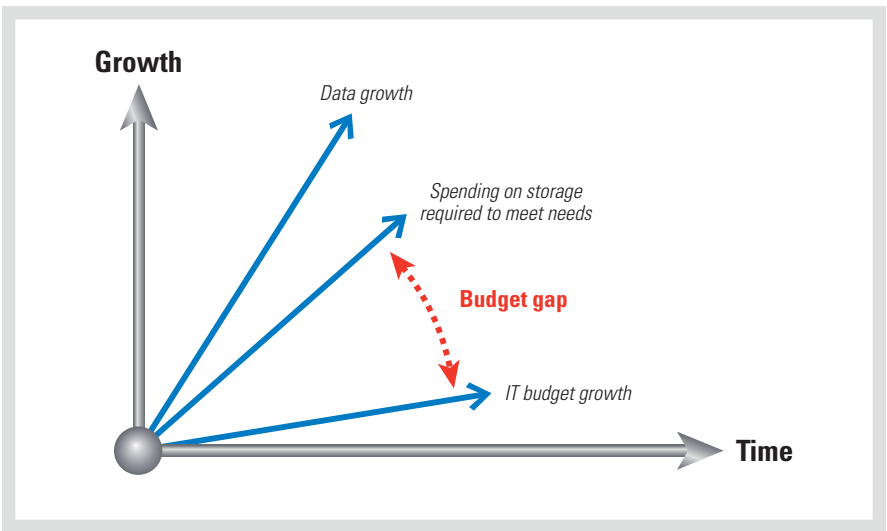


Figure 1. *Budget gap: Storage requirements are growing faster than typical IT spending plans*

[2] To learn more about the business benefits of IP-based networked storage in SMBs, see "iSCSI: Changing the Economics of Storage; Part 3—Using iSCSI in Small and Medium Businesses," by Travis Vigil, in *Dell Power Solutions,* November 2007, DELL.COM/downloads/global/power/ps4q07-20070402-Vigil.pdf.

[3] Support may vary by iSCSI array depending on vendor tests and certifications.

[4] To learn more about how iSCSI technology is being integrated into networked storage environments, see "iSCSI: Changing the Economics of Storage; Part 1—Understanding iSCSI in Enterprise Environments," by Travis Vigil, in *Dell Power Solutions,* May 2007, DELL.COM/downloads/global/power/ps2q07-20070335-Vigil.pdf.

In the meantime, a complementary mix of iSCSI and Fibre Channel provides a powerful and cost-effective approach, particularly for environments running a mix of applications. For example, iSCSI is well suited for applications with random I/O such as databases and virtualized servers, whereas Fibre Channel works well for high-throughput, low-latency applications with sequential I/O, such as streaming media and decision support software.

In high-bandwidth applications—especially those involving heavy transaction processing or high-speed, large-block data transfers—Fibre Channel can enable significant performance advantages, including functionality that can help increase utilization and simplify expansion for data-intensive applications. By combining iSCSI and Fibre Channel connectivity, organizations can capitalize on the benefits of both while moving toward network convergence on a simplified Ethernet infrastructure. For an example of how one company is taking advantage of both iSCSI and Fibre Channel storage, see the "Fluid Power Resource keeps data pumping with IP and Fibre Channel hybrid" sidebar in this article.

STORAGE-AWARE APPLICATIONS LEVERAGE DAS FOR COST-EFFICIENT GROWTH

Serial Attached SCSI (SAS), the successor to parallel SCSI, is designed to increase data transfer speeds and provide performance and reliability comparable to high-end systems. With broader reach and the ability to connect more devices per port than parallel SCSI, SAS offers a cost-effective way to optimize storage architectures.

In addition to enabling tiered direct attach storage (DAS), SAS is compatible with Serial ATA (SATA) and can run server drives as well as external and networked storage at exceptional speeds. Dell storage arrays incorporating SAS and SATA drives with iSCSI functionality for tiered DAS

HOW iSCSI STORAGE UNLOCKS THE POWER OF VIRTUAL SERVERS

iSCSI-based SAN technology allows applications to move around a virtualized data center environment while remaining attached to their storage. As a result, iSCSI is poised to take the business benefits of virtualization to the next level by helping simplify management and increase flexibility while increasing cost efficiency for the overall IT infrastructure.

For example, as a physical protocol that maps to a physical server, Fibre Channel requires administrators to manually rezone switches and change permissions when storage relationships change. iSCSI, in contrast, is an IP-based protocol that helps reduce the cost and complexity of deploying virtual servers compared with Fibre Channel by allowing administrators to map VMs directly to shared storage—in the same familiar way they manage the relationship between physical servers and shared storage (see Figure A).* This seamless VM mobility without complex reconfiguration allows unprecedented response to dynamic workload fluctuations and evolving business conditions. In addition, when an iSCSI network is configured with the proper logical or physical separations, iSCSI enables comparable security to Fibre Channel.

To create virtual servers with shared storage, organizations can deploy iSCSI alone or together with Fibre Channel in a tiered configuration. By offering products that support both iSCSI and a combination of iSCSI and Fibre Channel (see the "IP-based networked storage convergence changes economies of scale" section in this article), Dell enables organizations to optimize storage connectivity based on server workload.

For more information about Dell iSCSI storage, visit DELL.COM/iSCSI.

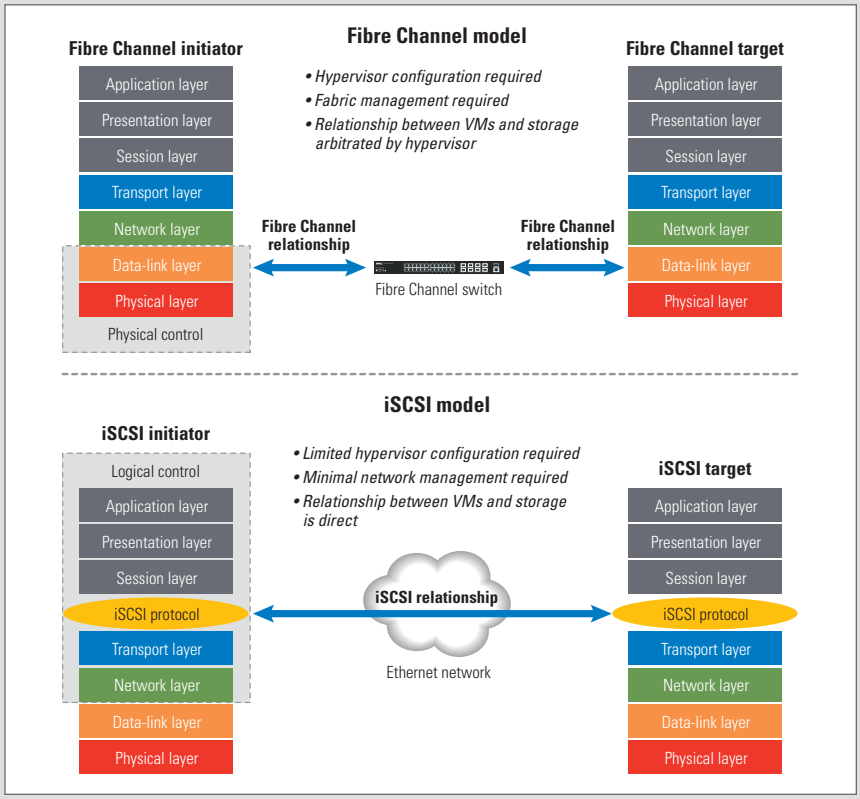


Figure A. *Unlike Fibre Channel, iSCSI allows administrators to map virtual machines directly to shared storage*

*For more information, see "iSCSI: Changing the Economics of Storage; Part 2—Deploying iSCSI in Virtualized Data Centers," by Matt Baker and Travis Vigil, in *Dell Power Solutions,* August 2007, DELL.COM/downloads/global/power/ps3q07-20070401-Baker.pdf.

applications include the Dell/EMC AX150i, Dell/EMC CX3 series, Dell PowerVault NX1950, and Dell PowerVault MD3000i.

As SAS and SATA drive options break through previous DAS scalability barriers with enhanced capacity and performance, many organizations are reconsidering DAS as a way to build out storage in cost-effective increments—particularly for burgeoning e-mail data.

Because e-mail is pervasive and protecting data generally costs more than simply storing it, the DAS model allows organizations to offer higher service levels and higher capacity with lower complexity and cost than a shared e-mail storage design. DAS helps simplify provisioning and data protection compared with shared networked storage because the rack of external hard drives in a DAS configuration is a logical extension of the host server's internal storage—and is managed the same way. In comparison, shared networked storage connected to multiple servers must be provisioned with security at the array level, which adds to deployment costs and management complexity.

Moreover, a number of critical enterprise applications—including Microsoft Exchange Server 2007, Microsoft SQL Server, and Oracle software—are being designed to facilitate storage-related tasks that enhance data availability, backup functionality, access management, and fault tolerance. Notably, applications with built-in replication capabilities are increasing the relevance of DAS deployments. Understanding how to take advantage of storage-related features in enterprise applications can also help businesses reduce storage

FLUID POWER RESOURCE KEEPS DATA PUMPING WITH IP AND FIBRE CHANNEL HYBRID

Fluid Power Resource, LLC, is an investment holding company that operates a premium group of 10 fluid power equipment distributors across the United States. The company employs approximately 1,100 people at 38 locations and is growing quickly. After a recent acquisition, the IT team decided to consolidate its distributed IT infrastructure in a central location. At the same time, the IT team needed a viable disaster recovery plan.

The distributed IT environment was configured to allow data in one location to access data from other locations should the power go out. However, downtime was common in the distributed IT environment, requiring substantial travel and expense to support ongoing operations and maintenance. To consolidate the IT environment and introduce the flexibility and availability needed for cost-efficient growth, the team deployed an EMC iSCSI/Fibre Channel storage platform with a blend of dedicated Dell PowerEdge™ physical servers running the Microsoft Windows Server® 2003 OS for some critical applications (such as Microsoft Exchange), and servers running VMware ESX Server virtualization software to host multiple VMs.

"We chose the EMC storage array for its ability to leverage both iSCSI and Fibre Channel connectivity simultaneously," says Jim Overly, director of IT at Fluid Power Resource. "This provides us with the flexibility to support all of our data access requirements in a single solution."

In addition, the array's simple storage management capabilities allow the IT team to manage information efficiently and exploit the capabilities of a mixed physical and virtual server environment. "The combination of Dell servers, VMware virtualization software, and the EMC storage array has provided our business with a comprehensive solution and ideal path to support future growth as business requirements evolve," Overly adds.

All critical information is now consolidated on highly available EMC storage arrays, enabling employees in multiple locations to access enterprise data when they need it. Most of the business applications access data through iSCSI; however, the IT team takes advantage of integrated Fibre Channel connectivity to support its Microsoft SQL Server database. For Overly, this translates into the performance users need and cost-effective scalability to support the company's growth and budget.

By simultaneously supporting iSCSI and Fibre Channel connectivity, EMC storage arrays enable Fluid Power Resource to support diverse data access requirements with a single, consolidated solution.

complexity and bridge the budget gap, especially when DAS may be a viable alternative or complement to networked storage arrays.

UNIFIED MANAGEMENT AND DATA PROTECTION REACHES FROM DESKTOP TO DATA CENTER

Many organizations are challenged to provide comprehensive data sharing, protection, and management—an urgent consideration for business continuity and compliance with regulations such as the Sarbanes-Oxley Act and the Health Insurance Portability and Accountability Act (HIPAA). The Dell Intelligent Data Management strategy is designed to enable a holistic environment of data protection and management. This approach entails capturing data at the point of creation; protecting it immediately, whether locally or remotely; ensuring integrity as diverse data types traverse multiple connections and travel through multiple systems; archiving, classifying, and searching; and ultimately deleting data at the end of the information life cycle.

A key cost-containment consideration is to match the expense of the storage platform to the value of the data. For example, bulk e-mail messages can be stored on inexpensive drives or even offline on tape, but sensitive materials that may be required for litigation or audits must be stored in a safe, secure way that allows quick and easy access. Open systems with intuitive, out-of-the-box storage and server management can allow organizations to back up, restore, and archive valuable intellectual property and transaction data in a cost-efficient, standards-based plug-and-play network environment.

STORAGE SIMPLIFICATION ALIGNS IT WITH STRATEGIC BUSINESS OBJECTIVES

As new voice and video technologies generate data by the terabyte, organizations have to change the economics of storage—not just to keep up with burgeoning capacity demands and regulatory requirements, but to free vital resources for business innovation and growth. Implementing a comprehensive storage strategy based on simple, cost-effective, and capable plug-and-play technologies can help reduce costs, increase efficiency, and ensure business continuity and compliance. A modular, open systems storage architecture based on custom pre-configured hardware and unified storage management enables organizations to streamline deployment and support, reduce complexity, and avoid the pitfalls of proprietary technologies that can impede business agility—all potential competitive differentiators in today's information-centric digital marketplace.

iSCSI-BASED SANs IN ACTION

To demonstrate the performance potential of an IP-based storage network in a consolidated storage environment, Dell engineers recently placed the application data for five Dell PowerEdge servers on one Dell PowerVault MD3000i array with two attached PowerVault MD1000 expansion enclosures. Located in the Austin, Texas–based Dell Enterprise Technology Center (DELL.COM/TechCenter), the team ran five separate workloads and performed two incremental backups over an eight-hour period to simulate a typical workday. In this test case, the results indicated that the PowerVault MD3000i effectively handled I/O throughput from the five example workloads—which included two nodes of a fully redundant Microsoft Exchange cluster, a Web-serving application, a Microsoft SQL Server database, and file-serving data transfers.*

In enterprise storage networks, the PowerVault MD3000i iSCSI array is designed to consolidate up to 16 fully redundant hosts, expanding to support up to 18 TB of data on SAS drives.

*For details, see "Storage Consolidation with the Dell PowerVault MD3000i iSCSI Array," by Dave Jaffe, Ph.D., and Kendra Matthews, in *Dell Power Solutions*, November 2007, DELL.COM/downloads/global/power/ps4q07-20080169-Jaffe.pdf.

Eric Endebrock *is a senior manager in the Dell Product Group, focusing on enterprise storage and data solutions. Eric has a dual B.S. in Management Information Systems and Accounting from the University of Nevada, Las Vegas, and an M.B.A. from Monash University.*

Dave Jaffe, Ph.D., *is a senior consultant on the Dell Enterprise Technology Center team who specializes in cross-platform solutions. He has a B.S. in Chemistry from Yale University and a Ph.D. in Chemistry from the University of California, San Diego.*

Kathryn White *is the features editor of* Dell Power Solutions *magazine. She has 26 years of development, communications, and marketing experience in the IT business. Kathryn has a B.S. in Mathematics from the University of South Carolina.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell storage solutions:
DELL.COM/Storage

Dell IT simplification:
DELL.COM/Simplify

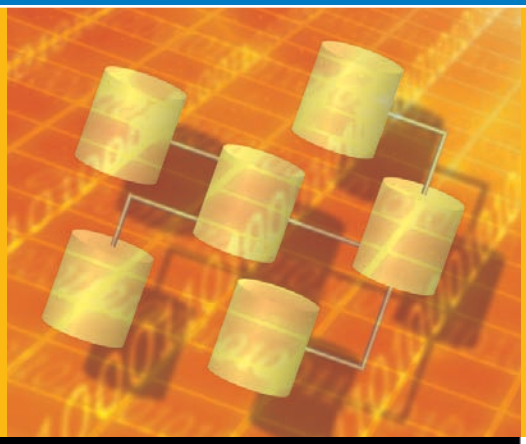
Dell Infrastructure Consulting Services:
DELL.COM/ICS

PAY UP TO 72% LESS FOR YOUR BUSINESS INTELLIGENCE

Dell and Microsoft can help reduce the cost and complexity of Business Intelligence. Our systems are quick to install, easy to manage, and built with standard components that work with what you already use — all for up to 72% less per terabyte than the competition.*

SIMPLIFY YOUR IT AT DELL.COM/SQL

*72% claim based upon a comparison of list prices of typical Business Intelligence offerings from leading hardware manufacturers versus Dell/Microsoft combined offerings. Benchmarked systems configured with 4-5 TB of data storage, database application software, and Business Intelligence analytic software. Dell is a trademark of Dell Inc.



iSCSI: CHANGING THE ECONOMICS OF STORAGE PART 3—USING iSCSI IN SMALL AND MEDIUM BUSINESSES

BY TRAVIS VIGIL

Small and medium businesses, remote offices, and departments and workgroups must cope with the same IT pressures as large enterprises, but with smaller budgets and fewer dedicated staff members. Internet SCSI (iSCSI)–based storage arrays such as the Dell™ PowerVault™ MD3000i offer a cost-effective way for these organizations to consolidate storage while providing powerful, easy-to-use management tools.

Related Categories:

Dell PowerVault storage

Dell/EMC storage

Internet SCSI (iSCSI)

Storage

Storage architecture

Storage area network (SAN)

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

Although small and medium businesses (SMBs), remote offices, and enterprise departments and workgroups have many of the same IT requirements as large enterprises, their limited resources can make it difficult to meet those requirements. As data growth rates continue to spiral upward, these organizations strive to meet their storage needs with limited budgets and small staffs, while simultaneously maintaining high levels of utilization and availability and managing robust backup and recovery processes. And because IT staff members may be performing multiple roles in these environments, simplified IT management can be critical to success.

Traditional attempts to meet scalability requirements within such constraints have typically included increasing storage capacity by adding servers and deploying additional tape drives. However, these approaches have their own disadvantages; in environments with single-application servers, for example, adding servers can result in poor storage utilization with some servers overflowing (such as e-mail systems) and others hardly utilized at all (such as Web servers). SMBs and similar organizations are constantly looking for technologies that can provide the capabilities they need while remaining both easy to manage and cost-effective. Internet SCSI (iSCSI)–based storage is designed to meet these requirements. By providing an entry point into storage area network (SAN) systems that allows these organizations to use standard, cost-effective Ethernet components rather than investing in a Fibre Channel infrastructure, iSCSI is well suited for their needs.

Past installments of this series have explored the basics of iSCSI, its advantages in enterprise environments, and how it can be deployed using Dell PowerVault and Dell/EMC storage,[1] as well as how data center administrators can combine iSCSI and virtualization to enhance efficiency and utilization.[2] This third installment focuses on how SMBs, remote offices, and enterprise departments and workgroups can use iSCSI to consolidate and simplify their storage environment, and how the Dell PowerVault MD3000i can help meet the needs of these organizations.

Understanding storage in small and medium businesses

Many SMBs, remote offices, and enterprise departments and workgroups have been relying principally on direct attach storage (DAS)—a simple and logical first step when expanding

[1] "iSCSI: Changing the Economics of Storage; Part 1—Understanding iSCSI in Enterprise Environments," by Travis Vigil, in *Dell Power Solutions*, May 2007, DELL.COM/downloads/global/power/ps2q07-20070335-Vigil.pdf.

[2] "iSCSI: Changing the Economics of Storage; Part 2—Deploying iSCSI in Virtualized Data Centers," by Matt Baker and Travis Vigil, in *Dell Power Solutions*, August 2007, DELL.COM/downloads/global/power/ps3q07-20070401-Baker.pdf.

capacity beyond servers' internal storage. However, this approach carries a number of disadvantages: it can only handle a limited number of attached hosts, can quickly become unwieldy to manage, and can require investment in additional disk- or tape-based storage systems as capacity requirements rise, resulting in overall poor capacity utilization.

Consolidating to a networked storage environment such as a SAN can help avoid these disadvantages, enabling many hosts to share resources and simplifying management. Organizations that find themselves working with multiple storage silos not optimized for their needs, having trouble scaling to accommodate data growth, or relying on burdensome backup and data protection processes based on multiple DAS systems are excellent candidates for storage consolidation. The same is true for remote offices and enterprise departments and workgroups that have been struggling to manage decentralized storage in different locations.

For these types of organizations, storage that is simple, capable, and cost-effective is key to success. In the past, they may have considered implementing a traditional Fibre Channel–based SAN to help increase utilization and simplify expansion. Fibre Channel does offer significant performance advantages in some environments, particularly those running applications with sequential I/O such as streaming media and decision support software. However, a Fibre Channel infrastructure also requires an investment in both hardware and training that may place it beyond the reach of organizations with limited resources.

An iSCSI-based SAN, in contrast, can provide the same advantages as a Fibre Channel–based SAN (including reduced total cost of ownership, increased capacity utilization, minimized backups, and simple manageability), but at a reduced acquisition cost. And iSCSI can provide high levels of performance comparable to Fibre Channel for many real-world applications common to SMBs, such as those with random I/O like Microsoft® Exchange and Microsoft SQL Server™ software. It offers comparable security to Fibre Channel when configured properly by logically or physically separating the iSCSI network. And its use of standard Ethernet components rather than specialized hardware makes iSCSI easy to manage for organizations with limited IT staff resources (see Figure 1).[3]

iSCSI also offers a high degree of flexibility, enabling administrators to integrate it into many different types of environments. The iSCSI protocol itself is widely supported by many common operating systems, applications, and other platforms, including Microsoft Windows® and Linux® as well as Microsoft Exchange, Microsoft SQL Server, Microsoft Cluster Service, Oracle® Real Application Clusters (RAC), and VMware® virtualization software.[4]

Consolidating through iSCSI-based storage can help simplify storage management while meeting the needs of SMBs and similar organizations in multiple ways. For SMBs implementing their first SAN, it can help them simplify management, reduce hardware and energy costs, and take advantage of common storage platforms for multiple applications without requiring a Fibre Channel implementation. For remote offices, departments, and workgroups in large enterprises, iSCSI can serve as a cost-effective second-tier SAN. In either case, it offers distinct advantages over both traditional DAS and Fibre Channel–based SANs, particularly for organizations constrained by limited staff and budget resources.

Introducing the Dell PowerVault MD3000i

By offering easy deployment, straightforward management tools, and powerful functionality in a cost-effective iSCSI-based storage system, the Dell PowerVault MD3000i is designed to meet the needs of SMBs, remote offices, and enterprise departments and workgroups looking to consolidate and simplify their storage environment. It is particularly well suited for those upgrading from DAS systems and implementing a SAN for the first time.

As an iSCSI-based storage system, the PowerVault MD3000i offers the benefits of consolidated SAN storage through a standard Ethernet infrastructure, without requiring the investment in hardware and expertise typically required when implementing a Fibre Channel–based infrastructure. Its advantages include the following:

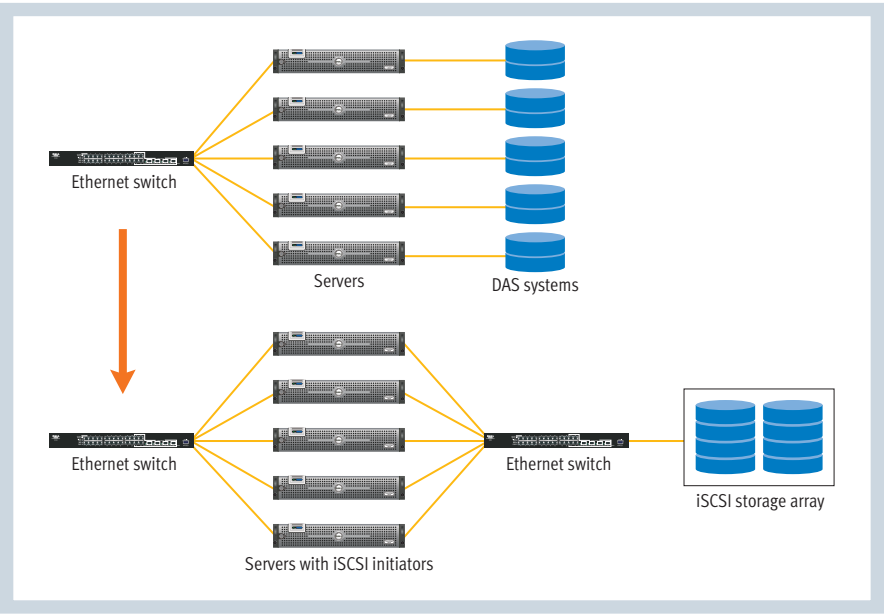


Figure 1. *Simplified storage through consolidation to an iSCSI-based SAN*

[3] For more information on the performance, security, and manageability of iSCSI, see "iSCSI: Changing the Economics of Storage; Part 1—Understanding iSCSI in Enterprise Environments," by Travis Vigil, in *Dell Power Solutions*, May 2007, DELL.COM/downloads/global/power/ps2q07-20070335-Vigil.pdf.

[4] Support may vary by iSCSI array depending on vendor tests and certifications.

- **Simplified deployment and management:** The PowerVault MD3000i is designed for easy deployment, and includes task-based Dell Modular Disk Storage Manager software to help administrators deploy and manage their storage.
- **High performance and availability:** The PowerVault MD3000i is designed to provide up to 400 MB/sec of throughput and includes dual active/active controllers, each with dual Gigabit Ethernet ports to avoid single points of failure.
- **Advanced data protection:** By supporting advanced functionality such as snapshot and virtual disk copy operations as add-on premium features, the PowerVault MD3000i can provide powerful, easy-to-manage data protection.
- **Easy scalability:** As data continues to grow, organizations can expand with up to 45 additional hard drives by attaching PowerVault MD1000 expansion enclosures.
- **Flexible configuration:** The PowerVault MD3000i supports 32- and 64-bit versions of Microsoft Windows Server® 2003 and Red Hat® Enterprise Linux 4 as well as Novell® SUSE® Linux Enterprise Server 9, and can be utilized in many different ways, including as part of server virtualization deployments, e-mail or database systems, and disaster recovery systems.
- **Cost-effective consolidation:** The PowerVault MD3000i is designed specifically for iSCSI, allowing organizations to utilize their existing Ethernet infrastructure to consolidate up to 16 fully redundant hosts.

Supported hosts	16 (fully redundant)
Supported expansion drives	45 (with PowerVault MD1000 expansion enclosures)
Total supported capacity	Up to 18 TB (using 400 GB Serial Attached SCSI [SAS] hard drives)
Mirrored cache size	Up to 1 GB per controller
Controllers	Dual active/active controllers
Ports	Two Gigabit Ethernet ports per controller
Expansion drive type	SAS
Supported RAID levels	RAID-0, RAID-1, RAID-5, and RAID-10
OS support	• 32- and 64-bit versions of Microsoft Windows Server 2003 Standard Edition and Enterprise Edition • 32- and 64-bit versions of Red Hat Enterprise Linux 4 Update 4 • Novell SUSE Linux Enterprise Server 9 Extended Memory 64 Technology (EM64T) with Service Pack 3 (SP3)

Figure 2. *Key features of the Dell PowerVault MD3000i*

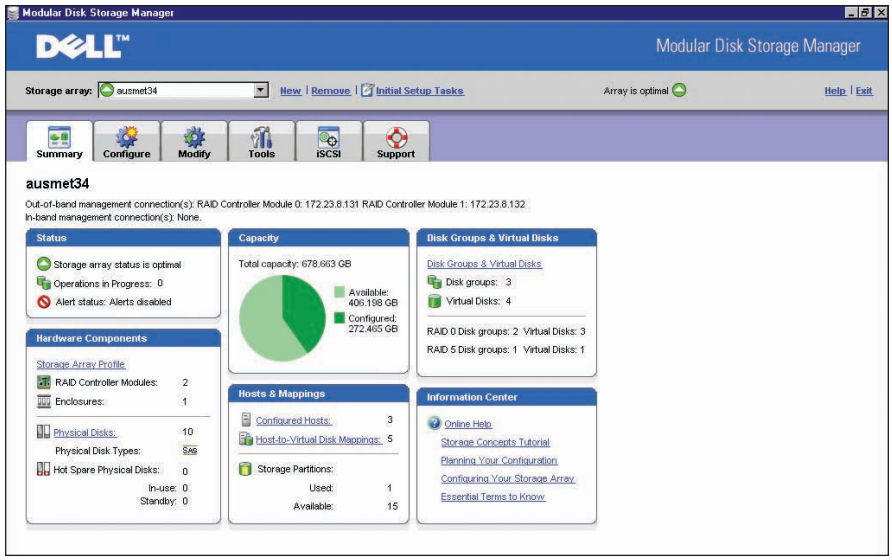


Figure 3. *Summary tab in Dell Modular Disk Storage Manager, providing array status and configuration information at a glance*

Figure 2 summarizes key features of the PowerVault MD3000i. Dell plans to expand both the feature set and OS support in the future, including support for Serial ATA (SATA) hard drives; the Microsoft Windows Server 2008 (code-named "Longhorn"), Red Hat Enterprise Linux 5, and Novell SUSE Linux Enterprise Server 10 operating systems; and the VMware ESX Server virtualization platform.

Simple, powerful management

Administrators can take advantage of multiple management tools with the PowerVault MD3000i, including the Dell Modular Disk Storage Manager graphical user interface (GUI), a command-line interface (CLI), Dell OpenManage™ IT Assistant, and the snapshot and copy functionality of Microsoft Volume Shadow Copy Service (VSS) and Microsoft Virtual Disk Service (VDS). Dell Modular Disk Storage Manager is designed for ease of use while offering powerful management functionality typically found in large enterprise systems. The CLI provides advanced functionality and scripting support. Integration with Dell OpenManage IT Assistant enables administrators to perform device discovery and monitoring through Simple Network Management Protocol (SNMP) as well as management tasks by launching Dell Modular Disk Storage Manager.

The primary GUI tool, Dell Modular Disk Storage Manager, integrates a task-based interface that administrators can use to view array status and other information and easily manage storage configurations and other settings (see Figure 3). Administrators can use these tools to

perform deployment tasks such as configuring host access, creating virtual disks and mapping them to hosts, and creating snapshots and virtual disk copies (if enabled); topology management tasks such as moving and renaming hosts and host ports and changing the host type; array-level tasks such as renaming an array and setting or changing passwords; iSCSI-specific tasks such as configuring Challenge Handshake Authentication Protocol (CHAP) and entering permissions for iSCSI initiators; and troubleshooting and support tasks such as viewing system event logs and downloading firmware.

Dell Modular Disk Storage Manager also includes the Recovery Guru feature designed to simplify troubleshooting. When problems arise, the Recovery Guru can provide a high-level description of the problem, additional information that helps identify the problem and where it resides, and step-by-step instructions to help administrators resolve the problem (see Figure 4). For administrators with many different responsibilities and limited time, this feature can be key to rapid problem resolution.

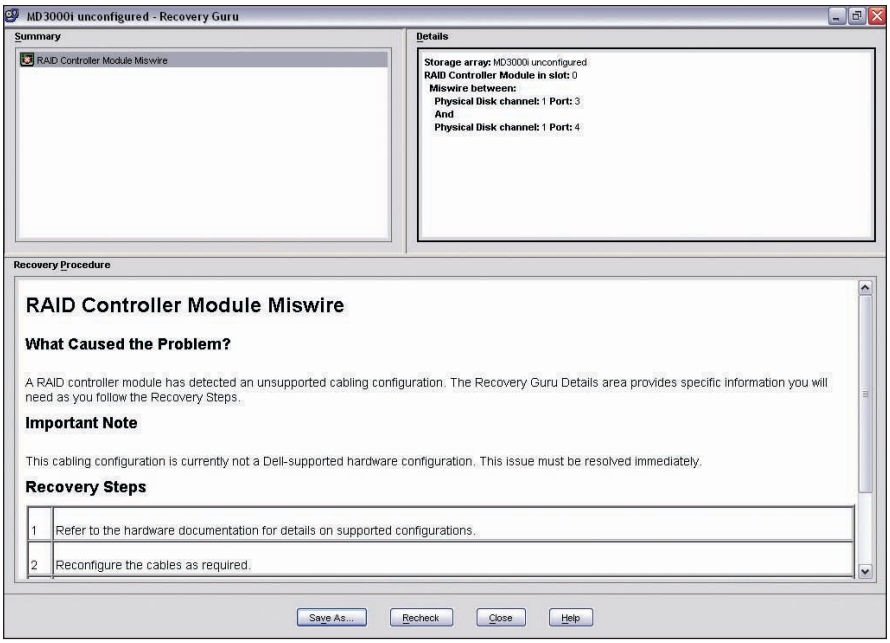


Figure 4. *Recovery Guru feature in Dell Modular Disk Storage Manager, providing troubleshooting details and step-by-step instructions to help resolve problems*

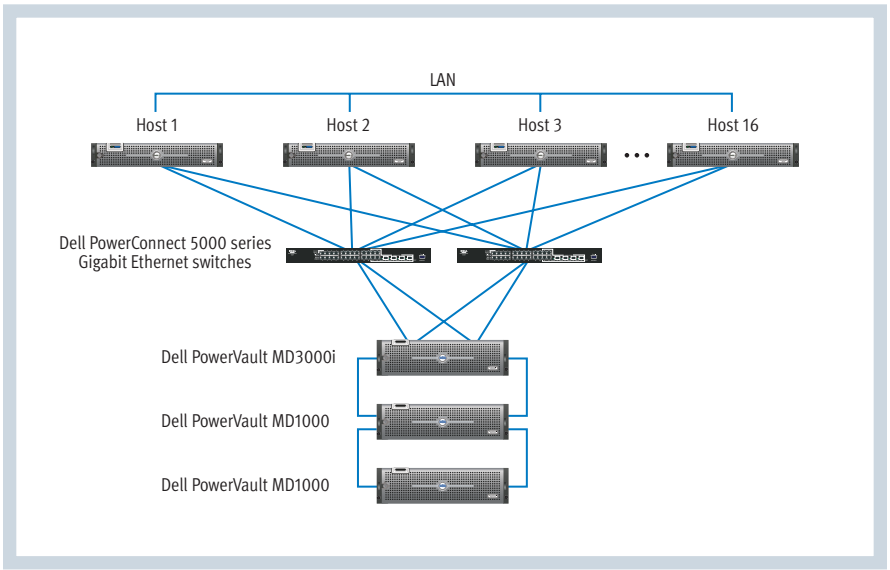


Figure 5. *Redundant multi-host configuration utilizing the Dell PowerVault MD3000i*

High availability and advanced data protection

Avoiding downtime and ensuring data availability are critical for many organizations, and SMBs are no exception. The PowerVault MD3000i is designed for high availability: by networking the hosts to the storage using iSCSI over standard Dell PowerConnect™ Gigabit Ethernet switches, administrators can create fully redundant paths for up to 16 hosts (see Figure 5). The redundant active/active controllers, automated I/O path protection with host-based multipath failover drivers, and automatic drive failure detection and rebuild functionality using global hot spare drives help ensure data availability even following a component failure. Redundant cooling and power systems offer an additional layer of protection.

Backup and recovery processes are also critical, helping protect against data loss following a system failure or disaster. For SMBs, making these processes as simple and efficient as possible for IT staff is often a key goal. Organizations can easily add virtual disk snapshot and copy functionality to Dell Modular Disk Storage Manager through premium features that support Microsoft VSS and VDS.

The VSS-based features of the PowerVault MD3000i are designed to allow organizations to create consistent, repeatable backup jobs, helping eliminate the need for manual coordination of applications, snapshots, and backup software—a task that can easily become burdensome for organizations with a limited IT staff. Integrating the PowerVault MD3000i with VSS-enabled applications enables administrators to easily create snapshots of data at a particular point in time, transport snapshots to a backup server, and perform other related tasks.

The VDS-based features are designed to provide consistent, simple administration for

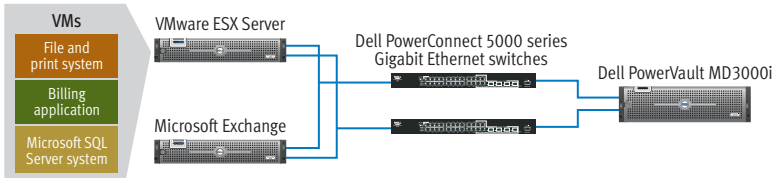


Figure 6. *Example deployment utilizing the Dell PowerVault MD3000i in a virtualized environment*

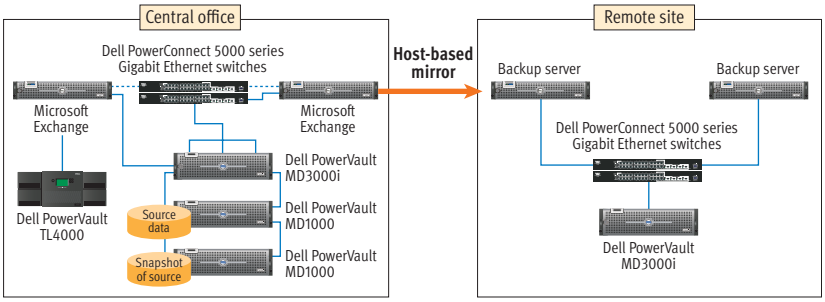


Figure 7. *Example deployment utilizing the Dell PowerVault MD3000i as part of a Microsoft Exchange system with host-based mirroring*

basic disk configuration tasks across heterogeneous storage platforms through a GUI or CLI. Administrators can use VDS to gain host access to the storage array, create RAID groups and virtual disks, and perform basic status monitoring.

Versatile support for different needs

The PowerVault MD3000i can provide iSCSI-based storage in a variety of environments, and can serve as a key part of virtualized data centers, e-mail or database systems, and disaster recovery systems. Figure 6, for example, shows an SMB deployment in which the file and print, billing, and Microsoft SQL Server systems run as virtual machines (VMs) on one server using the VMware ESX Server platform, with another server dedicated to Microsoft Exchange. Connecting these servers to the storage using iSCSI over Dell PowerConnect switches enables the PowerVault MD3000i to provide storage for all of these systems.

Figure 7 shows an example SMB deployment with the PowerVault MD3000i as part of a Microsoft Exchange e-mail system that also utilizes host-based mirroring to a remote site. In this environment, two attached PowerVault MD1000 expansion enclosures store the source data and snapshots of the source data. One Microsoft Exchange server stores backup data on a PowerVault TL4000 tape library for long-term archiving, while another uses host-based mirroring to replicate data to a remote site for disaster recovery. (Although host-based mirroring is typically slower than array-based mirroring, it is also typically more cost-effective, making it suitable for small deployments with limited budgets.)

Building capable, cost-effective iSCSI-based SANs

Although SMBs, remote offices, and enterprise departments and workgroups can benefit significantly from networked storage, limited staff and budget resources have often placed traditional Fibre Channel–based SANs out of reach. For these organizations, iSCSI may be the answer. By enabling networked storage over standard Ethernet components, iSCSI can offer advantages such as simplified management, efficient utilization, and easy scalability without requiring the investment in hardware and expertise typically required with Fibre Channel.

Designed with the needs of these organizations in mind—particularly those upgrading from DAS and implementing their first SAN—the Dell PowerVault MD3000i provides a simple, capable, cost-effective storage platform for iSCSI-based environments that can scale to meet data growth requirements. Its Dell Modular Disk Storage Manager tool combines an easy-to-use task-based interface with powerful functionality typically reserved for large enterprises, while features such as redundant controllers and optional integration with Microsoft VSS and VDS help protect against hardware failures and data loss. For SMBs and similar organizations, deploying iSCSI-based SANs with PowerVault MD3000i storage arrays can provide a cost-effective way to help ease management tasks, reduce ongoing hardware and energy costs, and build a flexible, consolidated storage environment.

Travis Vigil *is a product marketing strategist for Dell iSCSI and Dell PowerConnect solutions. He has nearly 10 years of experience with technology companies including Intel and Dell, and was most recently the product manager for Dell PowerVault disk storage. He has a B.S. from Stanford University and an M.B.A. from Northwestern University's Kellogg School of Management.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell iSCSI storage solutions:
DELL.COM/iSCSI

Dell PowerVault MD3000i:
DELL.COM/MD3000i

LEADER OF THE QUAD-CORE PACK.

THE COMPLETE LINE OF QUAD-CORE INTEL® XEON® PROCESSORS FOR MAINSTREAM SERVERS.

Intel continues to deliver the ultimate in business processing performance with its full range of Quad-Core Intel Xeon processors for servers. **Learn more at intel.com/xeon**

intel Xeon® inside™

STORAGE CONSOLIDATION WITH THE DELL POWERVAULT MD3000i iSCSI ARRAY

BY DAVE JAFFE, PH.D.
KENDRA MATTHEWS

The Dell™ PowerVault™ MD3000i Internet SCSI (iSCSI) array enables enterprises to easily consolidate storage for multiple servers. To illustrate its performance in a consolidated environment, Dell engineers performed tests demonstrating that the PowerVault MD3000i provides sufficient throughput to handle five server workloads as well as two incremental backups over an eight-hour period.

Related Categories:

Dell Enterprise Technology Center

Dell PowerVault storage

Internet SCSI (iSCSI)

Microsoft Exchange Server 2007

Storage

Storage consolidation

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

Enterprises continue to look for cost-effective ways to streamline their IT infrastructure and optimize data management. Although storage consolidation through the deployment of a storage area network (SAN) helps increase utilization and simplify management, in the past, this approach has typically required an investment in a Fibre Channel infrastructure. Today, Internet SCSI (iSCSI) offers an alternative by allowing the transmission of data packets over standard Ethernet networks, enabling enterprises to take advantage of existing networking expertise and equipment to help simplify SAN implementation and reduce barriers to wide-scale consolidation.

The Dell PowerVault MD3000i is an iSCSI-based modular disk storage array that can consolidate up to 16 hosts, expands to support up to 18 TB of data, and provides wizard-based installation, intuitive management, and advanced data protection software. Its modular expandability provides the flexibility to add capacity as needed: the storage array itself can house up to fifteen 3.5-inch Serial Attached SCSI (SAS) disk drives, and can easily be expanded to up to 45 drives by adding up to two PowerVault MD1000 expansion enclosures.

The array delivers a seamless suite of intuitive, intelligent storage management software capabilities. Dell Modular Disk Storage Manager can automatically configure the system for optimal performance and availability. The Recovery Guru tool can diagnose system problems and help determine appropriate steps for resolution. Optional snapshot and virtual disk copy features that support real-time backups of the entire data environment are designed to enhance data protection.

To help provide high availability, the PowerVault MD3000i supports redundant active/active controllers, management ports, and power and cooling systems designed to increase resiliency at the hardware level, and can automatically rebuild a failed drive using a global hot-spare drive.

To help demonstrate how the PowerVault MD3000i performs in a consolidated environment, in September 2007 Dell engineers placed the application data for five Dell PowerEdge™ servers on a single PowerVault MD3000i array with two PowerVault MD1000 expansion enclosures, then simulated a typical workday by running five separate workloads—two nodes of a fully redundant Microsoft® Exchange cluster, a Web-serving application, a Microsoft SQL Server™ database, and file-serving data transfers—during an eight-hour period. In addition, the test team performed two incremental backups using Symantec Backup Exec 11d for Windows Servers during this eight-hour simulation. As the results show, the PowerVault MD3000i provided sufficient throughput to handle the I/O from all of these workloads.

Dell PowerVault MD3000i features and test configuration

Figure 1 summarizes the key hardware features of the PowerVault MD3000i. The array includes one or two embedded

Available drives	15,000 rpm SAS drives available in 36 GB, 73 GB, 146 GB, or 300 GB capacities, and 10,000 rpm SAS drives available in 146 GB, 300 GB, or 400 GB capacities*
Rack size (including one enclosure with 15 disks)	3U
RAID cache size	512 MB per controller
Maximum number of disks (including two Dell PowerVault MD1000 expansion enclosures)	45
List price of tested configuration with three-year Gold-level support**	US$45,908

*For hard drives, GB means 1 billion bytes; actual capacity varies with preloaded material and operating environment and will be less.
**Price given as of October 4, 2007.

Figure 1. *Key hardware features of the Dell PowerVault MD3000i iSCSI storage array*

iSCSI controllers, each of which includes two Gigabit Ethernet iSCSI ports as well as an Ethernet management port and a SAS expansion port. In the Dell tests, the test team used a PowerVault MD3000i with two controllers and three enclosures containing a total of forty-five 146 GB, 15,000 rpm SAS drives.

Administrators can install Dell Modular Disk Storage Manager on a server connected to the same network as the array's Ethernet management port. They can then use this tool to create virtual disks—also known as storage logical units (LUNs)—and make them accessible to specific hosts. The Summary tab in Dell Modular Disk Storage Manager provides an overview of hosts and virtual disks as well as current storage array status (see Figure 2). The Tools tab allows administrators to set the array name and the IP address of the management port, and the iSCSI tab allows them to manage iSCSI settings such as the IP addresses of the four iSCSI ports, authentication, and target discovery (see Figure 3). The Configure Host Access (Manual) function in the Configure tab enables administrators to give a host server access to LUNs on the array, and the Edit Host Topology function in the Modify tab enables them to add the iSCSI initiator for each host (see Figure 4).

Administrators can create LUNs to hold host servers' application data either before or after adding those servers to the host topology. After creating a LUN, they can use the iSCSI initiator for the host OS (such as the Microsoft iSCSI Software Initiator used in the Dell tests) to configure a multipath connection to the LUN for load balancing and failover.[1] Figure 4, for example, shows the servers that had access to the PowerVault MD3000i during the Dell tests.

Figure 5 summarizes the disk configuration used in the Dell tests. In addition to creating the application data LUNs described in the preceding sections, the test team assigned disk 0 in enclosure 0 as a hot spare for the entire array and created a 10-disk RAID-10 snapshot repository on enclosures 1 and 2 to help protect against disk and enclosure failures.

Figure 6 illustrates the test configuration. The servers were connected to the PowerVault MD3000i array by two subnets, 10.10.20 and 10.10.22, using the Microsoft iSCSI Software Initiator to provide multiple connection paths for load balancing and failover. For simplicity, the PowerEdge 1950 also hosted the Dell Modular Disk Storage Manager management console for the PowerVault MD3000i. (Typically, the management server is on a separate network from the dedicated iSCSI networks.) Because the server running Symantec Backup Exec 11d created incremental backups from the passive Exchange node, it did not need to be connected directly to the PowerVault MD3000i.

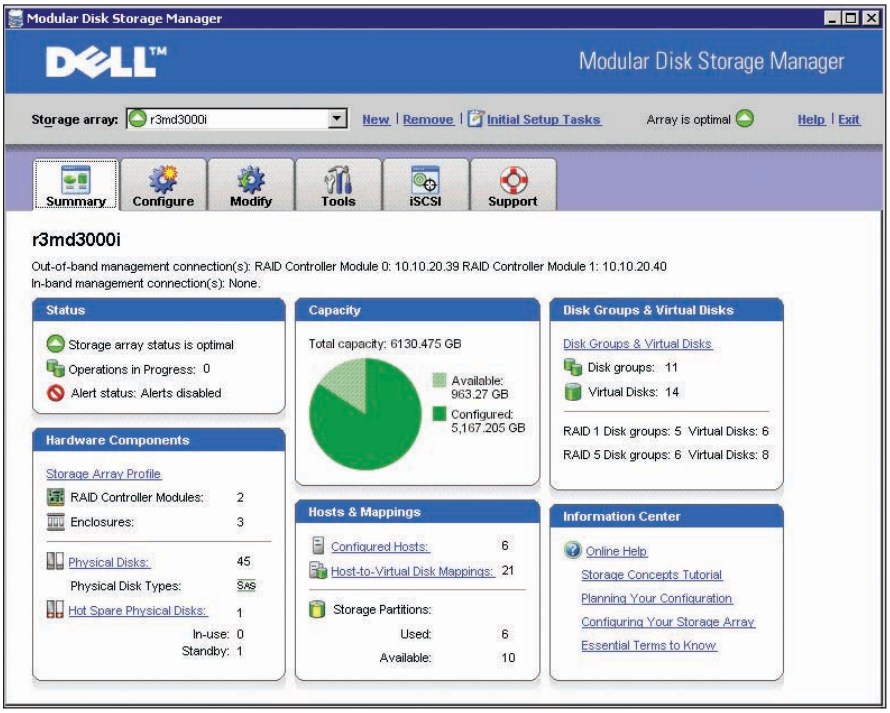


Figure 2. *Summary tab in Dell Modular Disk Storage Manager*

[1] For details on configuring a multipath connection on the Dell PowerVault MD3000i with Microsoft Multipath I/O, visit www.delltechcenter.com/page/MPIO+to+PowerVault+MD3000i+with+Microsoft+iSCSI+Initiator.

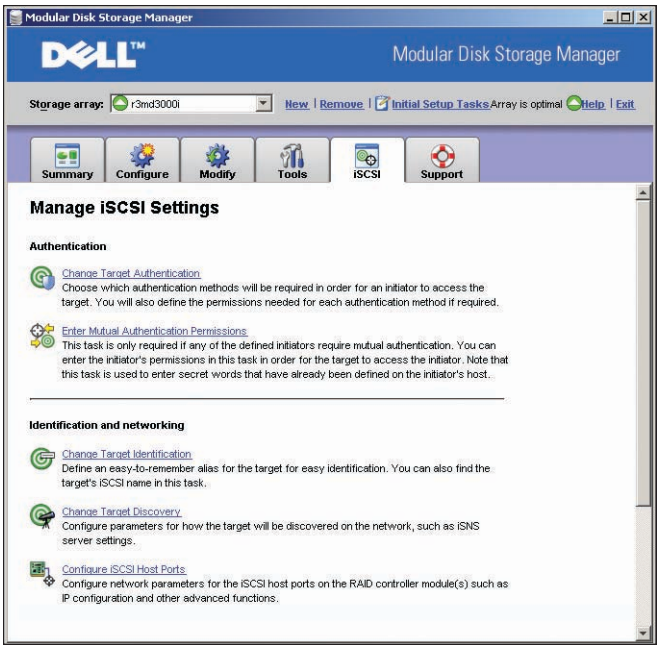


Figure 3. *iSCSI tab in Dell Modular Disk Storage Manager*

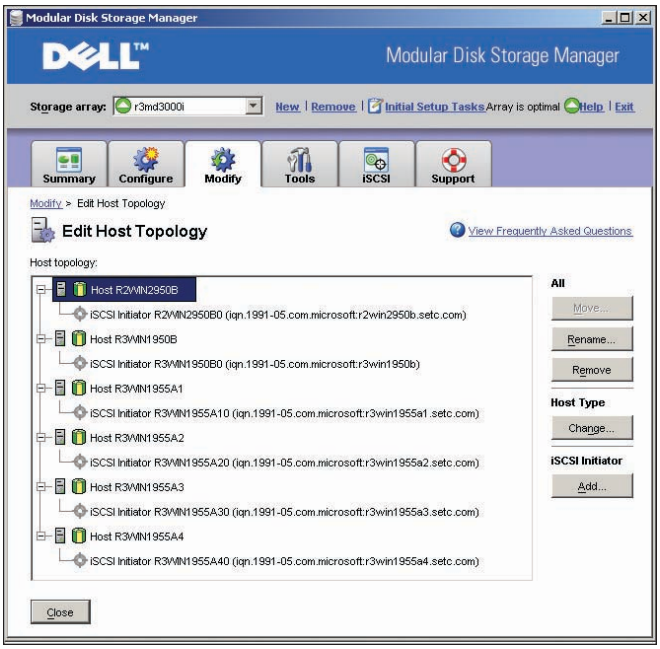


Figure 4. *Edit Host Topology screen in the Modify tab of Dell Modular Disk Storage Manager*

Test workloads

The Dell tests used five primary servers with application data consolidated on the PowerVault MD3000i: two servers running the Microsoft Exchange Server 2007 application, one acting as a Web server, one running the Microsoft SQL Server 2005 database platform, and one acting as a file server. A separate server running Symantec Backup Exec 11d functioned as a backup system for the Exchange database.

Microsoft Exchange Server 2007. The test team installed Exchange Server 2007 with cluster continuous replication (CCR) on two PowerEdge 1955 server blades, with one blade serving as the active node and the other serving as the passive node. CCR ships transaction logs from the active node to the passive node, helping keep the two nodes synchronized. If the active node fails, the Exchange users automatically fail over to the passive node. The test team created one two-disk RAID-1 log LUN and two four-disk RAID-5 data LUNs for the active node on enclosure 0, and a similar set of LUNs for the passive node on enclosure 1. (Full fault tolerance would require the passive node's LUNs to be on a separate storage array altogether.)

Enclosure	Disks					
	0	1–4		5–6	7–10	11–14
0	Hot spare	Available	Files (RAID-1)	Microsoft Exchange active log LUN (RAID-1)	Microsoft Exchange active data LUN 1 (RAID-5)	Microsoft Exchange active data LUN 2 (RAID-5)
1	Snapshot LUN (RAID-10)			Microsoft Exchange passive log LUN (RAID-1)	Microsoft Exchange passive data LUN 1 (RAID-5)	Microsoft Exchange passive data LUN 2 (RAID-5)
2	Snapshot LUN (RAID-10)			Microsoft SQL Server log LUN for Web server (RAID-1)	Microsoft SQL Server data LUN 1 for Web server (RAID-5)	Microsoft SQL Server data LUN 2 for Web server (RAID-5)
				Microsoft SQL Server log LUN (RAID-1)	Microsoft SQL Server data LUN 1 (RAID-5)	Microsoft SQL Server data LUN 2 (RAID-5)

Figure 5. *Disk configuration used in the test environment*

The test team used the Microsoft Exchange Load Generator (LoadGen) tool to simulate the Exchange workload. LoadGen simulates e-mail users and is typically run over an eight-hour period to simulate a complete workday. It comes with several default profiles; the test team used the heavy user profile, where each user averages 46 actions per day. LoadGen reports results in terms of average latency for each type of e-mail operation.

Web serving. To provide a Web-serving application, the test team used an ASP.NET–based Web application from the Dell DVD Store kit running under Microsoft Internet Information Services (IIS) in front of a SQL Server database, with both IIS and SQL Server running on a third PowerEdge 1955 server blade. The Dell DVD Store application provides a comprehensive set of login, search, and purchase screens for the simulated online store. The complete application code is freely available for public use

under the GNU General Public License (GPL) at linux.dell.com/dvdstore. The test team used a separate Web driver program from the DVD Store kit to simulate multiple users searching for and ordering DVDs from their Web browsers throughout the test. The back-end SQL Server database for this application used the large version of the DVD Store database (over 100 GB of data) and utilized one two-disk RAID-1 log LUN and two four-disk RAID-5 data LUNs.

Microsoft SQL Server 2005. The SQL Server 2005 workload also utilized the Dell DVD Store database, but without the Web interface. The test team installed SQL Server on a fourth PowerEdge 1955 server blade and built the large version of the DVD Store database on one two-disk RAID-1 log LUN and two four-disk RAID-5 data LUNs on the same disk group as the Dell DVD Store database used for the Web-serving application. Consequently, the same set of 10 disks was used simultaneously by two 100 GB databases during the eight-hour duration of the test. The test team used the SQL Server driver program included with the DVD Store kit to simulate a constant load of users logging in to the store; searching for DVDs by actor, title, or category; and purchasing DVDs.

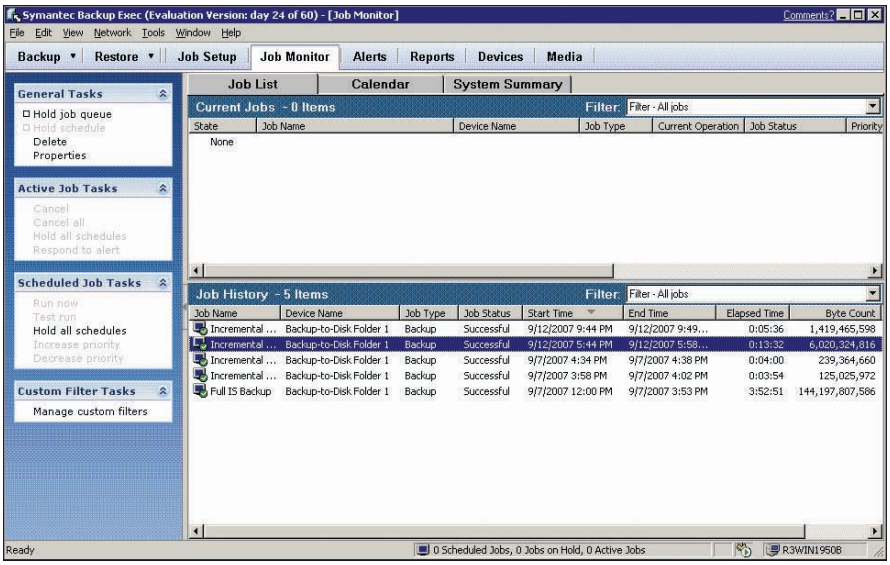


Figure 7. *Symantec Backup Exec 11d showing two incremental backups performed in the test environment*

File serving. To provide a file-serving workload, the test team created a two-disk RAID-1 LUN and assigned it to a PowerEdge 1950 server running the Microsoft Windows Server® 2003 OS. They assigned a drive letter to this LUN and shared it as a Windows Server 2003 file share, then copied a set of 41 files, with sizes varying from 1 MB to 3 GB, to it. They then mounted this file share on another PowerEdge 1950 server using the standard Map Network Drive function in Microsoft Windows® Explorer. To simulate use of this file share, the test team used a simple command-line shell script to copy files back and forth between this server and the file share. Two scripts ran in parallel during the eight-hour duration of the test, with delays placed in the scripts so that each copied approximately 40 files per hour.

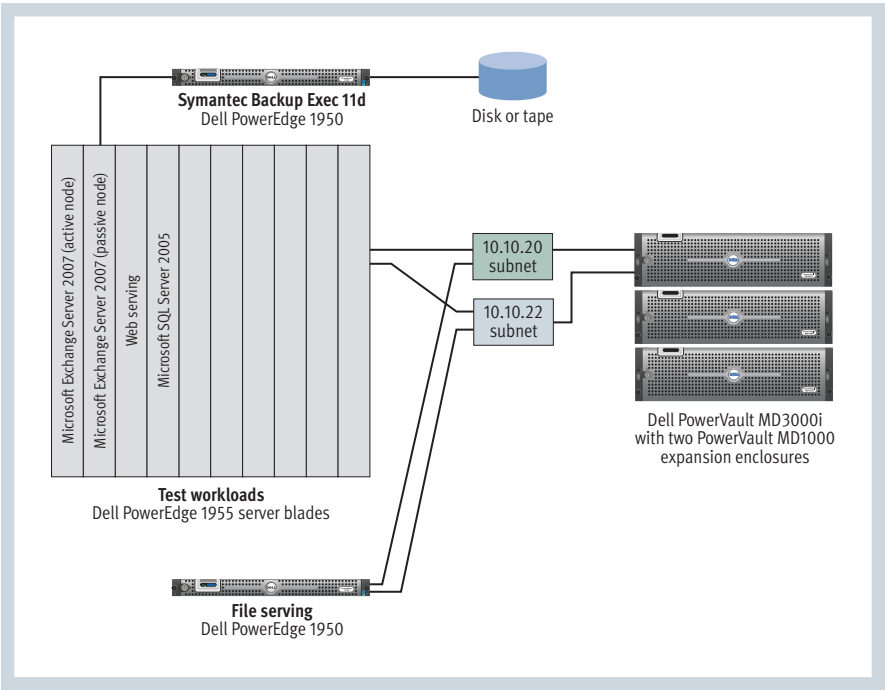


Figure 6. *Hardware configuration used in the test environment*

Symantec Backup Exec 11d. The test team used Symantec Backup Exec 11d, which enables fast, comprehensive backup and recovery using disk or tape, to back up the Exchange database to a PowerEdge 1950 server.[2] After Backup Exec 11d has performed a full backup of an Exchange database, it uses incremental backups from the passive node of the Exchange cluster to avoid affecting active Exchange users. The test team performed incremental backups with Backup Exec 11d at times corresponding to 11:30 A.M. and 3:30 P.M. during the simulated day (the two backups dated September 12, 2007, in Figure 7.)

[2] For more information, see "Protecting Critical Enterprise Applications with Symantec Backup Exec 11d," by Charles Butler, in *Dell Power Solutions*, May 2007, DELL.COM/downloads/global/power/ps2q07-20070408-Symantec.pdf.

Test results: Total work and data transfer

The test team ran the Microsoft Exchange Server 2007, Web serving, Microsoft SQL Server 2005, and file-serving workloads for eight hours to represent a workday running from 8 A.M. to 4 P.M., and performed two incremental backups with Symantec Backup Exec 11d as described in the preceding section. Each workload driver reported application-specific metrics, such as latency for Exchange tasks and orders per minute. The test team used Microsoft Windows Performance Monitor to record disk parameters.

Figure 8 summarizes the total work and data transfer of the five primary servers during the eight-hour test. The 500 LoadGen heavy users completed 23,832 tasks (such as reads, replies, and calendar updates) on the active Exchange node. The Exchange transaction log files were shipped continuously to the passive Exchange node, helping synchronize it with the active node. The passive node also handled the two incremental Backup Exec 11d backups. The Web-serving application handled 0.77 million orders, while the SQL Server database handled 1.1 million orders (equating to a total of more than 2 billion orders per year, if they received orders at this pace 24 hours a day). The file share copied 705 files back and forth, for a total transfer of 54 GB. Overall, the PowerVault MD3000i supplied 387 GB of data during the eight-hour test.

"By providing simple, cost-effective iSCSI storage along with a single easy-to-use management interface and optional integrated snapshot and virtual disk copy functionality, the PowerVault MD3000i enables enterprises to increase storage utilization by consolidating their storage with a single efficient system."

Simple, cost-effective iSCSI storage

As the Dell tests demonstrate, a Dell PowerVault MD3000i array with two attached PowerVault MD1000 enclosures was able to handle the I/O from five server workloads—two nodes of a fully redundant Microsoft Exchange cluster, a Web-serving application, a Microsoft SQL Server database, and file-serving data transfers—as well as two incremental backups with Symantec Backup Exec 11d. By providing simple, cost-effective iSCSI storage along with a single easy-to-use management interface and optional integrated snapshot and virtual disk copy functionality, the PowerVault MD3000i enables enterprises to increase storage utilization by consolidating their storage with a single efficient system.

Dave Jaffe, Ph.D., *is a senior consultant on the Dell Enterprise Technology Center team who specializes in cross-platform solutions. He has a B.S. in Chemistry from Yale University and a Ph.D. in Chemistry from the University of California, San Diego.*

Kendra Matthews *is a storage product marketing manager on the Dell Global Commercial Marketing team.*

	Total work	Total data transfer
Microsoft Exchange Server 2007 (active node)	23,832 tasks by 500 LoadGen heavy users	42 GB
Microsoft Exchange Server 2007 (passive node)	23,832 tasks by 500 LoadGen heavy users, as well as two incremental backups	82 GB
Web serving	0.77 million orders	61 GB
Microsoft SQL Server 2005	1.1 million orders	148 GB
File serving	705 file transfers	54 GB
Total		**387 GB**

Figure 8. *Total work and data transfer for the five primary workloads in the test environment*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINK

Dell PowerVault MD3000i:
DELL.COM/MD3000i

Network Nirvana

The Secure Choice for Open Convergence

Our integrated, open-architecture approach to understand and manage the priority and security of unified communications doesn't lock you into a particular voice, video or data vendor.

Let us show you how Enterasys Secure Networks™ solutions for switching, routing, wireless and intrusion detection/prevention can automatically discover, classify, prioritize and secure any convergence application without sacrificing performance.

It's VoIP without compromise as we support wired and wireless IP telephony solutions from Avaya, Cisco, Mitel, Nortel, Panasonic, Shoretel and Siemens, while also ensuring compliance with E-911 and CALEA regulations.

Leading companies in more than 70 countries ensure the integrity and performance of their IT services with Enterasys Secure Networks.™

We Secure *Any* Network

Download our **whitepaper on secure convergence** at www.enterasys.com/convergence to learn more about how we secure and prioritize a phone, camera, printer and computer all running on the same network. **When you're ready, call us at 877-801-7082 or arrange for a live demo** at www.enterasys.com/demo.

enterasys®
Secure Networks™

There is nothing more important than our customers.

THE ADVANTAGES OF IP-BASED NETWORKED STORAGE FOR MIDSIZE BUSINESSES

BY MIKE APIGIAN

Midsize businesses face the daunting challenge of managing rapid data growth with constrained budgets and limited staff. Advances in IP-based networked storage technology such as network attached storage and Internet SCSI (iSCSI) can help organizations of all sizes take advantage of existing resources to easily and cost-effectively build, maintain, and manage sophisticated storage networks.

Related Categories:

Dell/EMC storage

Internet Protocol (IP)

Internet SCSI (iSCSI)

Storage

Storage architecture

Storage area network (SAN)

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

Information is the fuel of today's enterprise, and in a global, 24/7 environment, it is more abundant then ever before. However, the infrastructure required to store, protect, and access large volumes of data efficiently and securely can be not only a valuable asset, but also a significant challenge.

Although the demand to consolidate data is felt at enterprises of all sizes, this need is particularly pronounced in midsize businesses. Restricted by limited budgets and staffing resources, IT managers at these organizations must still cope with rapid data growth and an increasing range of users, while simultaneously implementing new technologies and maintaining and enhancing existing ones. In these businesses, adding even one new administrator to help ease the burden can put an unacceptable strain on IT budgets. These businesses require technology that can not only help them overcome the growing challenges of data storage, but also allow them to take advantage of existing resources to do so.

Advances in IP-based networked storage technology such as network attached storage (NAS) and Internet SCSI (iSCSI) have created an opportunity for organizations of all sizes to cost-effectively build, maintain, and manage efficient, reliable storage networks using existing IP networks and familiar components. These sophisticated systems enable these organizations to achieve similar results as large enterprises, but in a simple, cost-effective way.

Understanding storage challenges in midsize businesses

Networked storage, which enables enterprises to consolidate and store massive amounts of data in a centralized storage system, can be a highly effective way to meet the challenges of rapid data growth. It can provide enhanced reliability, high availability, and high performance while helping ensure efficient and predictable data protection, backup and recovery, and long-term archiving.

Networked storage is very common in large enterprises, who have typically deployed high-bandwidth Fibre Channel–based storage area networks (SANs) as the foundation for critical large-scale applications. NAS offers a cost-effective supplement to dedicated Fibre Channel–based SANs, and is primarily suitable for file-sharing and file-serving applications. In combination, these two approaches can help create a powerful and reliable storage environment.

But while large enterprises typically have the resources to build and maintain these environments, small and midsize businesses may lack not only the budgets to invest in the technology, but also the staff resources to deploy and support it. Instead, they typically purchase and deploy storage as needed using either internal server storage or direct attach storage (DAS).

However, as the environment changes and data requirements grow, these separate servers and storage systems can

"Rapid data growth is both an asset and a challenge."

proliferate, creating a complex, inefficient, and difficult-to-manage environment with its own set of challenges:

- **Underutilization:** Distributed storage environments rarely maximize use of overall capacity, reducing return on investment in storage hardware.
- **Increased management costs:** The complexity of managing a large number of decentralized servers can quickly drive up management costs and overburden administrators.
- **Reduced reliability:** Spreading data and storage across the organization makes maintaining a reliable environment difficult, increasing the risk of data loss or unavailability and costly downtime.
- **Complex backup and recovery:** Because distributed storage does not offer a consistent, common, or simple way to perform backups, it can slow business continuity and disaster recovery processes—or render them ineffective.

Meeting storage challenges with IP-based networked storage

IP-based networked storage offers a simple, cost-effective way for midsize businesses to achieve the advantages of storage consolidation previously available primarily to large enterprises. And because many of these organizations already have an IP network in place, their administrators typically have the skills necessary to deploy and manage IP-based storage.

IP-based storage offers multiple advantages over the distributed storage common to midsize businesses, including the following:

- **Increased utilization:** Consolidated IP-based storage enables servers to access and share storage, helping maximize utilization of these resources.
- **Reduced management costs:** Consolidated storage enables centralized management, helping simplify administrative tasks and reduce management costs.
- **Increased reliability:** A shared set of dedicated IP-based storage systems can help significantly increase the reliability and availability of application data.
- **Simplified backup and recovery:** IP-based networked storage enables administrators to easily implement consistent, common, and simple backup and recovery processes.

When planning and deploying an IP-based networked storage environment, midsize businesses may have questions about costs, ease of transition, application support, and the capacity for growth. For more information on such factors, see the "Considerations when transitioning to IP-based networked storage" sidebar in this article.

Building simple, cost-effective storage environments

Rapid data growth is both an asset and a challenge, and for midsize businesses, meeting that challenge with limited budgets and IT staff can be difficult. By deploying IP-based networked storage such as NAS and iSCSI, these organizations can gain the same IT advantages as large enterprises—including increased utilization, reduced management costs, increased reliability, and simplified backup and recovery—in a simple, cost-effective way.

Mike Apigian *is a manager of marketing programs at EMC. Since joining EMC in 1998, he has had responsibilities for storage platforms and software offerings. Mike has a B.S. in Business Administration from Norwich University and a High Tech M.B.A. from Northeastern University.*

CONSIDERATIONS WHEN TRANSITIONING TO IP-BASED NETWORKED STORAGE

Compared with internal server storage or DAS, IP-based networked storage offers a highly efficient, cost-effective way to handle rapidly increasing data. But before planning and deploying this type of storage, midsize businesses may have questions such as the following:

- **Does deployment require a major budgetary investment?** Because IP-based storage can utilize existing networks and cost-effective IP connectivity, it is typically even within the reach of organizations with limited budgets.
- **Will the transition be difficult?** IT staff at midsize businesses are often already handling a wide range of administrative duties and other IT initiatives, making simple deployment a critical factor. IP-based networked storage utilizes technology typically already familiar to IT staff, helping provide a seamless, simple transition.
- **What types of applications can it support?** Effective storage must support a wide range of applications. NAS is well suited for file sharing and file serving, while iSCSI is well suited for e-mail and online transaction processing applications, providing sufficient performance for typical applications with small-block, random I/O loads.
- **Is it flexible enough to support changing needs?** Storage requirements can change as organizations evolve. When choosing an IP-based storage system, organizations should take into account not only the initial deployment and consolidation, but also the system's scalability and additional functionality that may be required in the future. In planning their storage environments, midsize businesses may want to include the flexibility to offer IP and Fibre Channel to support a variety of application requirements, as well as the ability to support advanced backup, recovery, and archiving systems.

INTELLIGENT E-MAIL ARCHIVING: CLASSIFICATION, FILTERING, RETENTION, AND DISCOVERY

BY ART GILLILAND

Implementing an e-mail archiving system can bring multiple challenges, including choosing which messages to archive, how long to retain them, and how to find specific messages after they have been archived. The Symantec® intelligent archiving approach enables organizations to efficiently classify, filter, retain, and search for e-mail messages while helping simplify management and control resource costs.

Related Categories:

Backup, recovery, and archiving (BURA)

Microsoft Exchange Server 2007

Regulatory compliance

Storage

Symantec

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

With the recognition that e-mail has become a critical part of IT infrastructure, many organizations are reevaluating their e-mail management policies and systems. Across many industries and public-sector organizations, IT professionals must address three common concerns regarding e-mail: resource management, retention management, and discovery management.

Given these challenges, enterprises across the world are evaluating or using software-based e-mail archiving systems. These systems are typically designed to help IT staff control e-mail storage costs while giving end users simplified storage and search functionality and providing legal departments with a consistent system for retaining and finding e-mail messages across the enterprise. As IT departments plan or implement these systems, however, they must account for several important considerations:

- **Storage size:** Although e-mail archives can provide rapid return on investment, they also create a high demand for storage. And because they may have to retain this data for many years, IT departments are seeking ways to optimize their archival storage costs.
- **Retention period:** E-mail archiving can force a necessary but challenging discussion within organizations about how long they should keep e-mail messages. Many enterprises and government bodies have retention policies for traditional paper records, yet struggle to determine the appropriate policies for e-mail.
- **Search functionality:** Finally, organizations must estimate the amount of data they will accumulate in their archives over time and look for ways to reduce the search time and costs for finding the data they need.

Although archiving systems help greatly simplify the problems with resources, retention, and discovery that plague many e-mail environments, they do not eliminate them. Many of the key challenges that remain stem from the fact that although e-mail messages may share fundamental characteristics—a sender, a recipient, a subject, and a body—they do not all have the same value. For example, Figure 1 shows two e-mail messages from a company CEO. The e-mail on the left is a critical company document, an official record that may drive a series of business actions to help this company compete—and that may serve as evidence in the future if these actions are investigated for being anticompetitive. In contrast, the e-mail on the right is important to the CEO but not to the company's future (unless, perhaps, his son is the head of BETA Corporation). Yet many e-mail archiving environments treat messages as though they all have the same value.

To help overcome this limitation, enterprises should look for ways to intelligently classify, filter, retain, and search for

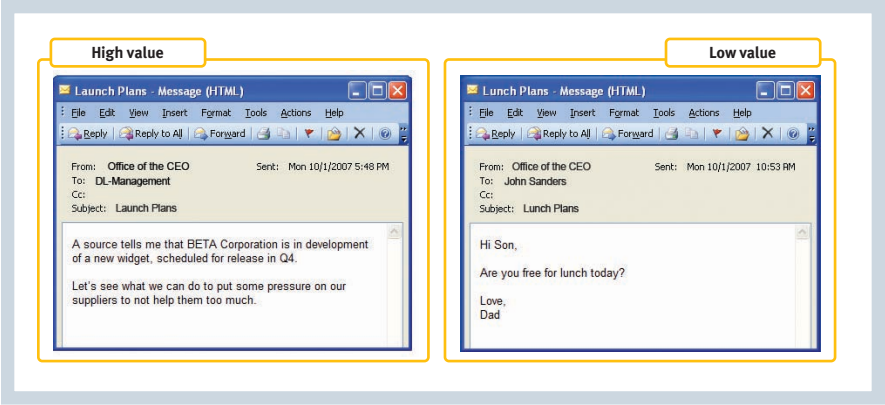


Figure 1. *E-mail messages with significantly different enterprise value*

messages, enabling them to handle messages differently based on their value, subject, and other considerations. The Symantec intelligent archiving approach can help enterprises implement this type of system and create an efficient automated system for e-mail archiving and retrieval.

Understanding e-mail archiving and the Symantec intelligent archiving approach

E-mail archiving systems typically work by first capturing e-mail messages from the environment, either immediately (referred to as journaling) or after a period of time such as 30 days. They then store those messages for a period of time defined by the administrator, referred to as the retention period. Finally, these systems index the messages, their properties, and their attachments so that legal, finance, human resources, and other groups can find them later.

When implementing an e-mail archiving system, organizations must make three fundamental policy decisions: what they should archive, how long they should retain archived messages and related data, and how they can find messages and data later. Deciding how long to retain information is perhaps the most challenging of these three. On the one hand, many enterprise leaders would like to keep e-mail messages as long as possible: e-mail is vital to enterprise operations, and many employees frequently go back to old messages for information. On the other hand, legal and IT professionals often see the downside of retaining e-mail. First, every additional message retained increases storage and IT costs. Second, keeping some messages longer than necessary may increase risk for the organization later on. Finally, as the number of archived e-mail messages grows, it becomes difficult to locate individual messages when needed.

Among organizations that use e-mail archiving systems, many of them use systems that fall into one of three groups:

- **Non-automated archiving system:** Organizations that lack an automated system still archive e-mail messages—but typically do so in a way that is often inefficient, ineffective, and risky. IT staff may archive e-mail messages by retaining e-mail server backups, while management may discover that messages that were deleted from the e-mail server years ago still remain on a backup tape or notebook computer. These unwanted backups may cause problems if the organization is forced to turn over data it did not even know it possessed to an opposing litigant or investigator.
- **Automated archiving system that keeps everything for the same period of time:** IT departments have driven many early e-mail archiving deployments to help reduce e-mail storage costs and increase application efficiency. Many organizations have archiving systems that retain all e-mail for the same period of time, be it one year, three years, five years, or more. Many of these organizations have not yet reached the point where they must actively expire e-mail; however, those that have reached the end of their retention period often extend it, just to be safe.
- **Automated archiving system that keeps everything forever:** Some early adopters of e-mail archiving based their implementation on regulatory mandates. Because these mandates were often vague in scope and length of time, some organizations have indefinite retention policies for their archives as they await further clarification from the government or depend on other organizations to take the first step.

The Symantec intelligent archiving approach is a natural evolution of early software-based e-mail archiving systems (see Figure 2). It is designed to provide an automated archiving system that treats e-mail messages differently based on value and enterprise requirements while simultaneously providing simplified search and retrieval functionality through the following components:

- **Intelligent classification:** Deciding which e-mail messages are relevant for which purposes

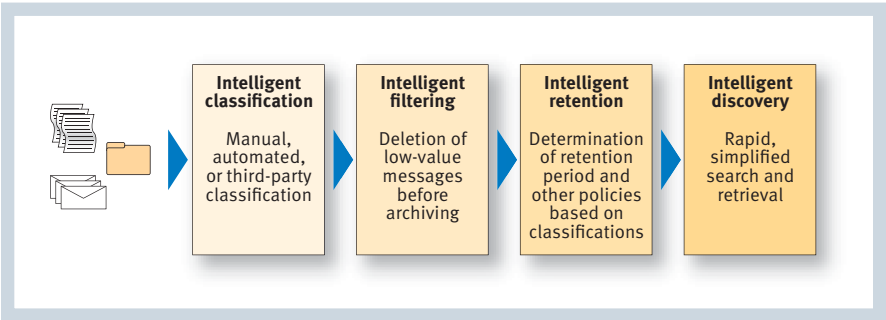


Figure 2. *Key components of the Symantec intelligent archiving approach*

- **Intelligent filtering:** Discarding irrelevant e-mail messages before archiving, helping reduce the size of the archive
- **Intelligent retention:** Determining how long to keep archived e-mail messages based on their classification
- **Intelligent discovery:** Tagging e-mail messages with metadata during archiving, helping simplify future retrieval

Taking the first step: Intelligent classification

Intelligent classification is critical to the Symantec intelligent archiving approach, enabling organizations to overcome key challenges of classification and differentiate the enormous number of e-mail messages they send and receive each day.

Key classification challenges

E-mail can be classified into a number of different categories, ranging from the very basic ("Business" or "Personal") to the very sophisticated ("2005 Tax Records," "Reseller Contracts for Germany," and so on). This type of classification is not new—records management has existed for over a century, with enterprises and government organizations devoting substantial time, money, and personnel to storing official documents in files, placing those files in boxes, storing those boxes in warehouses, and keeping track of the entire process. However, e-mail has introduced three challenges that make previous records management approaches inadequate: volume, universality, and informality.

Volume. Many organizations receive huge quantities of e-mail messages. In traditional records management models, because organizations might deal with thousands of official records, the threshold for creating a record was very high. Someone had to print or write a document, submit it to a records clerk (or have it be part of a defined process), and so on. Now, all it takes to create a record is for someone to click the Send button.

Essentially, e-mail can happen at nearly the speed of thought, rather than at the speed of print—and the volume increases accordingly. An organization with 10,000 end users sending and receiving 100 messages per day with 200 working days per year would create 200 million messages per year. Over a five-year period, that amounts to a billion messages.

Universality. Under traditional records management processes, records were typically created by defined groups—legal, finance, human resources, and other departments—that could be trained on organizational policies based on compliance demands. Now, anyone in an organization across the globe might be sending e-mail messages, and contractors and outsourcing partners can increase the complexity of this problem even further.

With myriad individuals across countries, languages, time zones, and enterprise boundaries, organizations are challenged to disseminate and enforce documented e-mail retention policies and guidelines. Many organizations are not willing to stake their reputation and financial security on trusting every user to follow the process.

Informality. Informality is perhaps the trickiest problem of the three. E-mail messages, unlike memos or faxes, are notoriously informal. A thread about last weekend's activities can quickly transform into a discussion about this quarter's sales forecast, and a casual comment, when taken in proper context, can take on major importance.

Approaches to intelligent classification

There are three primary methods of intelligently classifying e-mail: manual (having end users classify their messages), automated (having an archiving system classify messages), and third-party (having another system classify messages, either manually or automatically). Figure 3 summarizes the advantages and disadvantages of each.

Manual classification. Although a key advantage of archiving is avoiding the need for end users to make classification decisions, many organizations have concluded that blending automated archiving with some human oversight is necessary. In this approach, a user lets the archiving system know how to classify an e-mail message in the archive from within the organization's e-mail software, such as the Microsoft® Office Outlook® client. One method involves presenting a folder structure defined by the IT department to the user in Outlook (in addition to his or her normal personal folders). For example, a salesperson using Outlook might see the folders shown in Figure 4.

The advantage of manual classification is that end users can often judge a message's value more accurately than an automated algorithm. At the same time, however, this approach increases work for end users and can lead to inaccuracies from user error or malicious intent.

Automated classification. The opposite approach is to rely on the archiving system to classify messages. For many organizations, an ideal automated classification engine could identify what each message is and its relevance to the enterprise.

	Description	Advantages	Disadvantages
Manual classification	End users classify messages by dragging them into a folder or by selecting options from a pop-up window	May provide better judgment of message values than an automated algorithm	Requires additional work for end users, and end users may inadvertently or deliberately misclassify messages
Automated classification	The archiving system classifies messages by analyzing message properties and content	Uses a consistent, repeatable process with a low burden on end users	Has potential for false categorizations
Third-party classification	A third-party records management or gateway system classifies messages	Can take advantage of existing classification systems and handle multiple types of content	Same as for manual and automated systems, depending on the underlying approach

Figure 3. *Comparison of intelligent classification methods*

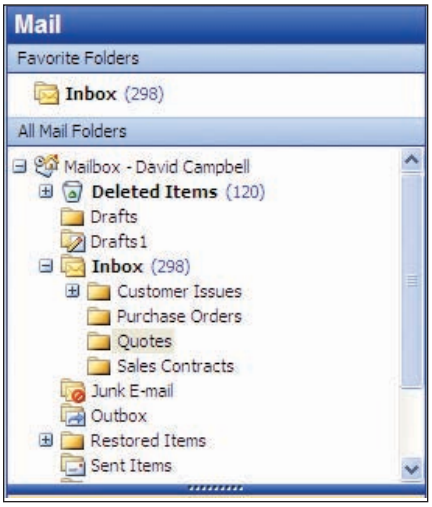


Figure 4. *Example folders displayed in Microsoft Office Outlook*

Classification engines typically use a combination of approaches to analyze a message and determine the type of content, including using information such as the following:

- **Senders and recipients:** Messages from the legal department, for example, typically contain legal content.
- **Keywords or phrases:** Messages and attachments with a "confidential" disclaimer, for example, typically identify data that could be stored as intellectual property.
- **Patterns:** Messages containing numbers in the form ###-##-####, for example, typically contain Social Security numbers, and might be classified as patient information for a hospital and require special retention rules.

In contrast to manual classification, the automated approach places a limited burden on end users and decreases the risk of data being misclassified from user error or malicious intent. However, like other automated systems, classification systems are subject to false categorizations.

Third-party classification. Many organizations are deploying systems to categorize and manage records across multiple content types. They can then integrate these systems with leading e-mail archiving systems to allow the archive to store and optimize e-mail messages while enabling the records management system to drive retention decisions that are consistent across different types of data.

Putting classification to work: Filtering, retention, and discovery

After sorting through e-mail messages and classifying them, Symantec intelligent archiving can take three actions: filtering, retention, and discovery.

Intelligent filtering

For many organizations, not everything needs to be archived—messages sent to all employees may not require archiving for every mailbox, and personal e-mail may not require archiving at all. Filtering out these noncritical messages helps reduce the total cost of ownership of archiving systems.

Intelligent retention

Intelligent retention bases retention policies on classifications the organization has established. For example, records managers can define a set of categories that map to distinct retention periods. The system can then determine the appropriate retention period for each message based on these categories. This approach helps reduce the risk of keeping some messages too long while not retaining others long enough.

Intelligent discovery

Classification systems can tag messages with metadata to enable effective search and retrieval. Some organizations review e-mail on a daily basis and may want to filter out messages that are clearly personal. Others may want to tag e-mail messages from the legal department as "possibly privileged" to help reduce the time required for future searches.

Implementing intelligent archiving with Symantec Enterprise Vault

When implementing an archiving system, each IT department must assess its own objectives and requirements and decide on an appropriate approach. But regardless of the direction they ultimately choose, many organizations would be well served to add intelligence to their archiving policies.

The Symantec Enterprise Vault™ application provides a software-based intelligent archiving platform to store, manage, and search for enterprise data from e-mail systems, file server environments, instant messaging platforms, and content management and collaboration systems. Because not all data has equal value, Enterprise Vault utilizes intelligent classification and retention technologies to capture, categorize, index, and store target data to help enforce policies, protect enterprise assets, reduce storage costs, and simplify management. It also integrates specialized applications such as Discovery Accelerator and Compliance Accelerator to help enterprises mine archived data in support of legal discovery, content compliance, knowledge management, and information security initiatives.

In addition, Dell and Symantec have partnered to deliver a comprehensive Secure Exchange solution designed to protect critical Microsoft Exchange environments. Featuring state-of-the-art components from Dell and Symantec, this end-to-end solution is based on a modular reference architecture independently validated by Symantec for high performance, flexibility, and scalability. Enterprises can combine this solution with Symantec Enterprise Vault to facilitate intelligent e-mail archiving and regulatory compliance in Exchange environments.

Deploying a robust, efficient e-mail archiving system

Symantec intelligent archiving enables organizations to efficiently classify, filter, retain, and search for e-mail messages. By implementing intelligent archiving with Symantec Enterprise Vault, enterprises can create a robust e-mail archiving system to help simplify management and control resource costs while meeting regulatory and enterprise requirements.

Art Gilliland *is vice president of product management at Symantec.*

ENTERPRISE WORKLOADS FOR SMALL AND MEDIUM BUSINESSES ON THE DELL POWERVAULT NX1950

BY DAVE JAFFE, PH.D.
TODD MUIRHEAD

The Dell™ PowerVault™ NX1950 integrated network attached storage system offers cost-effective, versatile storage for small and medium businesses and remote offices. To demonstrate its performance in a high-availability configuration, Dell engineers simulated a mixed environment running key Microsoft® Exchange and SQL Server™ workloads alongside Microsoft Windows® and Linux® file shares.

Related Categories:

Dell PowerVault storage

Microsoft Exchange Server 2007

Microsoft SQL Server 2005

Network attached storage (NAS)

Performance characterization

Storage

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

The Dell PowerVault NX1950 integrated network attached storage (NAS) system offers a unified way for small and medium businesses (SMBs) and remote offices to consolidate both block application data and file data in a single storage device. It provides block application data for e-mail, database, and other enterprise applications through Internet SCSI (iSCSI), and file data through both Common Internet File System (CIFS) on Microsoft Windows operating systems and Network File System (NFS) on Linux, UNIX®, and Mac OS operating systems.

iSCSI allows communication between servers and storage through standard Ethernet network interface cards (NICs), switches, and cables, providing a simple, cost-effective way to implement storage area networks (SANs). In addition to deploying storage for enterprise applications, many organizations provide access to data through CIFS and NFS file shares residing on NAS. By providing both iSCSI storage logical units (LUNs) as well as CIFS and NFS file shares, the PowerVault NX1950 provides the advantages of both a SAN and NAS in a unified system.

To evaluate how the PowerVault NX1950 could perform in a typical SMB environment, in July 2007 Dell engineers performed tests in which the PowerVault NX1950 served as storage for a simulated 1,000-employee online sales company. In the test environment, Microsoft Exchange Server 2007 and Microsoft SQL Server 2005 provided employee e-mail and a back-end e-commerce database, respectively, using iSCSI LUNs on the PowerVault NX1950. At the same time, simulated employees accessed both CIFS and NFS file shares through Windows and Linux.

Dell PowerVault NX1950 features and test configuration

The PowerVault NX1950 integrated high-availability configuration consists of two 1U, two-socket controllers running the Microsoft Windows Unified Data Storage Server 2003 OS and providing redundancy capabilities through Microsoft Cluster Service. The controllers connect through Serial Attached SCSI (SAS) to a PowerVault MD3000 disk array with redundant embedded RAID controllers containing a total of 1 GB cache. Figure 1 summarizes the key hardware features of this configuration.

In addition to integrating a PowerVault MD3000 disk array, the PowerVault NX1950 can be expanded with up to two PowerVault MD1000 disk expansion enclosures for a total of 45 drives. The Dell tests used a PowerVault NX1950 high-availability configuration with three enclosures containing forty-five 146 GB, 15,000 rpm SAS drives.

Windows Unified Data Storage Server 2003 provides an integrated console for storage system management that includes tasks for disk, LUN, and user management, as well as tools for creating and maintaining iSCSI targets and both

> *"By providing both iSCSI storage LUNs as well as CIFS and NFS file shares, the PowerVault NX1950 provides the advantages of both a SAN and NAS in a unified system."*

Available drives	15,000 rpm SAS drives available in 36 GB, 73 GB, 146 GB, or 300 GB capacities, and 10,000 rpm SAS drives available in 146 GB, 300 GB, or 400 GB capacities*
Rack size (including Dell PowerVault MD3000)	5U
RAID cache size	1 GB
Maximum number of disks (including two Dell PowerVault MD1000 expansion enclosures)	45
List price of tested configuration with three-year Gold-level support and remote installation**	US$70,987

*For hard drives, GB means 1 billion bytes; actual capacity varies with preloaded material and operating environment and will be less.
**Price given as of July 30, 2007.

Figure 1. *Key hardware features of the Dell PowerVault NX1950 integrated high-availability configuration*

CIFS and NFS file shares. The console view in Figure 2 shows the configuration used in the Dell tests. The test team created LUNs for CIFS (FileLUN1) and NFS (NFS_LUN) files, as well as data and log LUNs for Microsoft Exchange and SQL Server. To help maximize flexibility and security, the team created separate iSCSI targets for SQL Server and Exchange (R1NXISCSI and R1NXISCSI2, respectively).

Figure 3 summarizes the disk configuration used in the Dell tests. Enclosure 0 included a four-disk RAID-5 LUN for CIFS files, a five-disk RAID-5 LUN for general iSCSI use, five available disks, and a global hot-spare disk for the entire array. Enclosure 1 contained the Exchange LUNs: a four-disk RAID-10 log LUN and two five-disk RAID-5 data LUNs. Enclosure 2 contained the SQL Server LUNs: a two-disk RAID-1 log LUN and two five-disk RAID-5 data LUNs. Enclosure 2 also contained a two-disk RAID-1 LUN for NFS files. Finally, two mirrored RAID-1 disks containing the cluster quorum disk were placed on enclosures 1 and 2 to help protect against disk and enclosure failures.

The test team installed Exchange Server 2007 on a Dell PowerEdge™ 2950 server and connected this server through iSCSI to the R1NXISCSI2 target on the PowerVault NX1950. They added the iSCSI Qualified Name (IQN) of this server to this target on the PowerVault NX1950, which provided access to the Exchange LUNs listed in Figure 3, with the log LUN mounted as drive L:\ and the data LUNs mounted as drives M:\ and O:\. They installed SQL Server 2005 on another PowerEdge 2950 server and connected this server through iSCSI to the R1NXISCSI target on the PowerVault NX1950. They added the IQN of this server to the PowerVault NX1950, which provided access to the SQL Server LUNs listed in Figure 3, with the log LUN mounted as drive I:\ and the data LUNs mounted as drives G:\ and H:\.

Figure 4 illustrates the test configuration. The PowerEdge 2950 servers running Exchange Server 2007 and SQL Server 2005 were connected to the PowerVault NX1950 array by two subnets, 10.10.20 and 10.10.22, using the Microsoft iSCSI Software Initiator to provide multiple connection paths to the array for load balancing and failover. The servers simulating the file users attached only through the 10.10.22 subnet.

The test team designated the four services—two iSCSI targets, CIFS, and NFS—on the PowerVault NX1950 to run in active mode on the first node of the PowerVault NX1950 management server cluster and in passive mode on the second node. To help maximize performance,

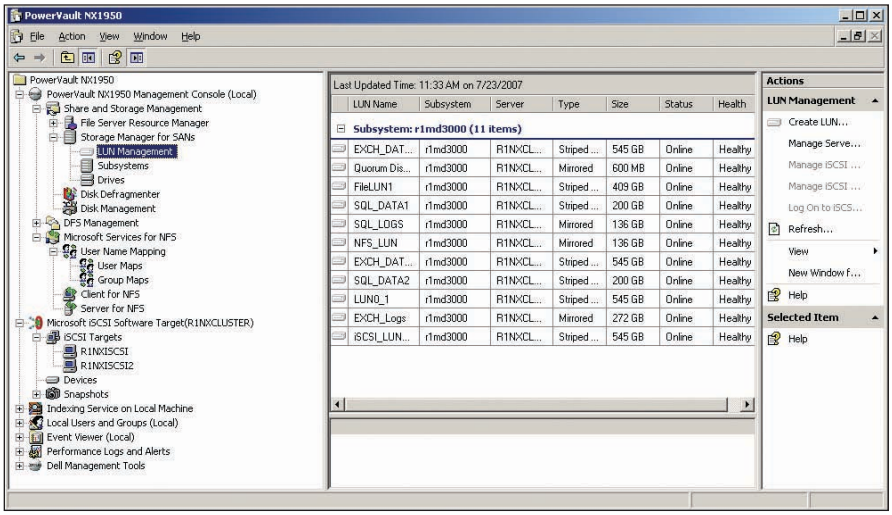


Figure 2. *Microsoft Windows Unified Data Storage Server 2003 console for the Dell PowerVault NX1950*

Enclosure	Disks			
	0	1–4	5–9	10–14
0	Hot spare	CIFS LUN (RAID-5)	iSCSI LUN (RAID-5)	Available
1	Quorum (RAID-1)	Microsoft Exchange log LUN (RAID-10)	Microsoft Exchange data LUN 1 (RAID-5)	Microsoft Exchange data LUN 2 (RAID-5)
2	Quorum (RAID-1)	NFS LUN (RAID-1) / Microsoft SQL Server log LUN (RAID-1)	Microsoft SQL Server data LUN 1 (RAID-5)	Microsoft SQL Server data LUN 2 (RAID-5)

Figure 3. *Disk configuration used in the test environment*

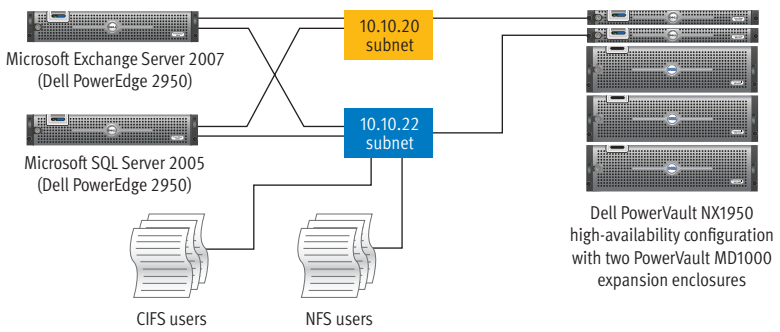


Figure 4. *Hardware configuration used in the test environment*

services can be load balanced between two servers by running two in active mode on each server; however, if a node fails, all services would need to run on a single node. Therefore, to help demonstrate that a single node of the PowerVault NX1950 cluster can handle the full load, the test team set all services to active mode on the first node.

Test workloads

The test team used four workloads as part of an eight-hour simulation: the Microsoft Exchange Load Generator (LoadGen) tool, the Dell DVD Store database workload, a CIFS file-serving workload, and an NFS file-serving workload.

Microsoft Exchange Load Generator. The test team used LoadGen, a test tool provided by Microsoft, to simulate 1,000 heavy e-mail users running against the Exchange Server 2007 server. LoadGen is typically run over an eight-hour period to simulate a complete workday, and includes several default profiles. The heavy user profile used in the Dell tests resulted in users averaging 46 actions (reads, replies, calendar updates, and so on) per day. LoadGen reports results in terms of average latency for each type of e-mail operation.

Dell DVD Store. The Dell DVD Store database simulates the back end of an online e-commerce application; the complete application code is freely available for public use under the GNU General Public License (GPL) at linux.dell.com/dvdstore. The test team built the large version of the DVD Store database—over 100 GB of data—on the SQL Server LUNs listed in Figure 3. They used the SQL Server driver program included with the DVD Store kit to simulate a constant load of users logging in to the store; searching for DVDs by actor, title, or category; and purchasing DVDs during the eight-hour duration of the test.

CIFS file serving. For the CIFS file-serving workload, the test team stored a set of 41 files, with sizes varying from 1 MB to 3 GB, on the CIFS LUN on the PowerVault NX1950 and on a PowerEdge 2950 server running the Microsoft Windows Server® 2003 OS and simulating multiple desktops. They then mounted the PowerVault NX1950 CIFS file share on the PowerEdge 2950 using the Map Network Drive tool in Windows Explorer, and copied the files back and forth using a simple Windows command shell script. Two scripts ran in parallel, with delays placed in the scripts so that each script copied approximately 40 files per hour.

NFS file serving. The test team simulated the NFS file-serving workload similarly to the CIFS workload. They copied the same set of 41 files to the NFS LUN on the PowerVault NX1950 and to a PowerEdge 2950 server running the Red Hat® Enterprise Linux 4 OS. Using the NFS client on this server, they mounted the PowerVault NX1950 NFS file share to the Linux file system. A simple Korn shell script copied the files back and forth, with approximately 40 files copied per hour.

"As these Dell tests demonstrate, a single PowerVault NX1950 with 45 attached drives can help meet the e-mail, database, and file-serving requirements of enterprises with up to 1,000 employees."

Test results: Total work, throughput, and data transfer

Figure 5 summarizes the results of the Dell tests. The 1,000 simulated heavy Exchange users ran without error during the simulated eight-hour day while providing response times of less than a second. The SQL Server database received 5,500 orders per minute for a total of 2.6 million orders during the eight-hour day (equating to almost 3 billion orders per year, if it received orders at this pace 24 hours a day). The simulated file users copied a total of 133 GB of data back and forth on the CIFS and NFS file shares, for an average of 133 MB per simulated employee. In total, these four applications required the PowerVault NX1950 to handle 13.3 MB/sec of throughput, transferring a total of 373 GB during the eight-hour day.

"As a unified solution, the PowerVault NX1950 can provide cost-effective, versatile, easy-to-use storage for SMBs."

Figure 6 shows the PowerVault NX1950 throughput curve during the test in detail. The average throughput was 13 MB/sec (13,876,412 bytes/sec, as shown in Figure 6), with a peak of 132 MB/sec (138,216,742 bytes/sec, as shown in Figure 6).

	Total work	Throughput	Total data transferred
Microsoft Exchange Server 2007	1,000 users with 46,081 e-mail operations performed	4.1 MB/sec	116 GB
Microsoft SQL Server 2005	2.6 million orders submitted	4.4 MB/sec	124 GB
CIFS file share	650 files copied	2.2 MB/sec	61 GB
NFS file share	674 files copied	2.6 MB/sec	72 GB
Total		**13.3 MB/sec**	**373 GB**

Figure 5. *Dell PowerVault NX1950 throughput and data transfer test results*

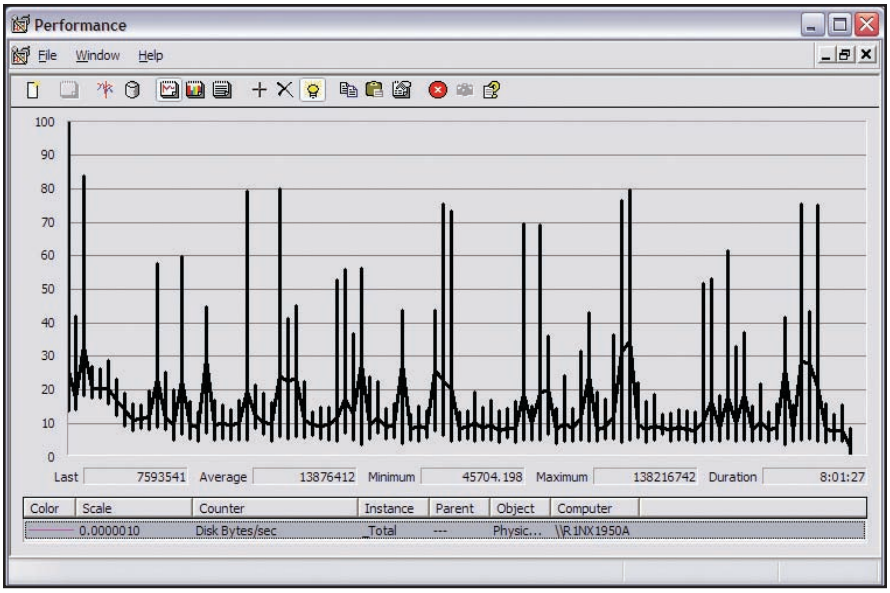


Figure 6. *Dell PowerVault NX1950 throughput during the eight-hour test, including the average, minimum, and maximum throughput*

Cost-effective, versatile, easy-to-use storage

The Dell PowerVault NX1950 integrated NAS solution is well suited for SMBs needing to share files and looking to deploy an iSCSI-based SAN for their key application data. As these Dell tests demonstrate, a single PowerVault NX1950 with 45 attached drives can help meet the e-mail, database, and file-serving requirements of enterprises with up to 1,000 employees. As a unified solution, the PowerVault NX1950 can provide cost-effective, versatile, easy-to-use storage for these organizations.

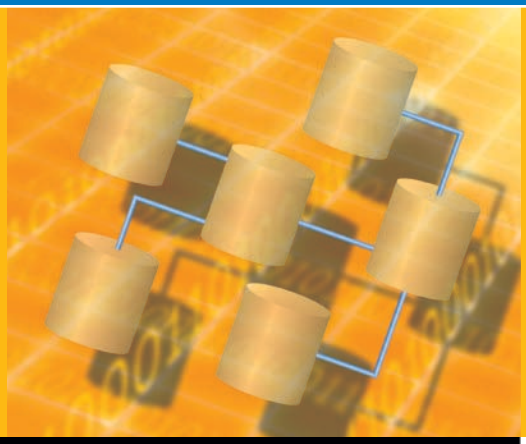
Dave Jaffe, Ph.D., *is a senior consultant on the Dell Enterprise Technology Center team specializing in cross-platform solutions. He has a B.S. in Chemistry from Yale University and a Ph.D. in Chemistry from the University of California, San Diego.*

Todd Muirhead *is a senior engineering consultant on the Dell Enterprise Technology Center team, where he works with database, messaging, virtualization, and storage solutions. Todd has a B.A. in Computer Science from the University of North Texas.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINK

Dell PowerVault NX1950:
DELL.COM/NX



CONFIGURING MICROSOFT SQL SERVER 2005 ON DELL SERVERS AND STORAGE

BY WENDY CHEN
SUMNER STEWART
PAUL RUBIN
CHETHAN KUMAR
ALVIN PENG

Following best practices from Dell and Microsoft can help enterprises design a high-performance Microsoft® SQL Server™ 2005 database on Dell™ PowerEdge™ servers and Dell PowerVault™ storage. This article provides specific guidance and best-practice recommendations in the areas of storage, OS, and database configuration.

Related Categories:

Database

Dell PowerEdge servers

Dell PowerVault storage

Microsoft SQL Server 2005

Storage

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

Applying best practices derived from laboratory and real-world experiences can help IT professionals design and configure reliable, high-performance database environments. Microsoft and Dell offer several best practices for implementing, configuring, and managing Microsoft SQL Server 2005 database solutions in enterprise architectures that include Dell PowerEdge servers, Dell PowerVault MD3000 Serial Attached SCSI (SAS) storage enclosures, and the Microsoft Windows Server® 2003 OS. Dell PowerEdge servers are designed to deliver high performance for critical enterprise applications such as databases, messaging, Web services, and infrastructure. The Dell PowerVault MD3000 is a high-performance storage array built for critical applications running on one, two, or a pair of clustered PowerEdge servers. By combining Dell PowerEdge servers with Dell PowerVault MD3000 storage systems, administrators can easily deploy PowerEdge servers as building blocks of database solutions. Working in tandem, Dell servers and storage can help consolidate and virtualize computing and storage resources for SQL Server–based environments.

Figure 1 illustrates an architecture for SQL Server 2005 on Dell hardware. This architecture comprises the following components:

- **Client systems:** The clients access data stored within the SQL Server 2005 database.
- **Client-server network:** The connecting network consists of network controllers, cables, and switches.
- **Dell PowerEdge servers:** The servers run Windows Server 2003 and SQL Server 2005.
- **Server-storage interconnect:** The servers and storage devices connect with SAS.
- **Dell PowerVault storage:** The storage hardware supports up to 45 drives.

Preparing the storage systems

The physical disks in the Dell PowerVault MD3000 storage array provide physical storage capacity for the SQL Server 2005 database. Before data can be stored, the PowerVault MD3000 physical storage capacity must be configured into components known as *disk groups* and *virtual disks*. A disk group is a set of physical disks that are logically grouped and assigned a RAID level. Virtual disks are logical entities that the server uses to store data. Each disk group provides the overall capacity needed to create one or more virtual disks.

SQL Server 2005 uses three specific storage areas for tempdb, the transaction log file, and data for user-defined databases. Tempdb is a shared working area for all databases on the server and for activities such as creating temporary tables, sorting, and processing sub-queries. The transaction log file stores the details of modifications made

> *"Working in tandem, Dell servers and storage can help consolidate and virtualize computing and storage resources for SQL Server–based environments."*

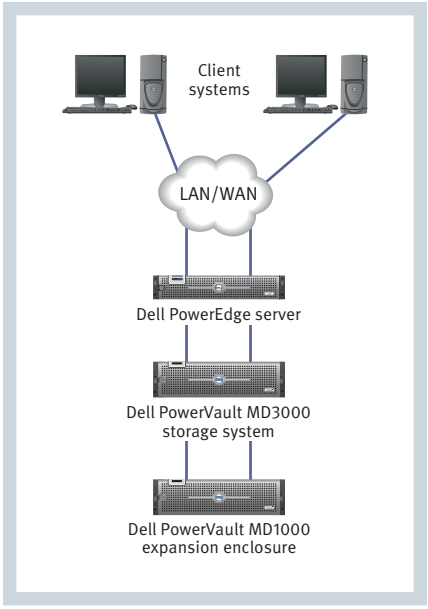


Figure 1. *Architectural overview for Microsoft SQL Server 2005 on Dell hardware*

to a SQL Server database along with the details of the transactions that performed each modification. Data for user-defined databases is stored in a separate data storage area. This data file contains user-defined database objects such as tables, indexes, and stored procedures.

Microsoft best practices recommend separating the tempdb and transaction log files onto separate physical disks—which, for the Dell PowerVault MD3000, means separating these files onto their own virtual disks on separate disk groups. This separation can enhance I/O performance by ensuring that these files do not share the same physical disks. Figure 2 illustrates a sample disk group and virtual disk configuration that separates tempdb, transaction logs, and data files onto distinct physical disks.

Creating disk groups and virtual disks

Microsoft recommends two common best practices with respect to storage configuration for SQL Server 2005. The first best practice is to spread database objects across multiple disk spindles; the second is to limit the number of data files based on the number of processor cores.[1] For design simplicity and to help avoid the complexity introduced by having too many virtual disks, the sample configurations described in this article also map the number of virtual disks for data to the number of processor sockets. For example, a two-socket system (with either dual-core or quad-core processors) would have two virtual disks for data. Figure 3 illustrates sample virtual disk configurations and sizes for four-socket host servers.

RAID-10 is considered the optimal choice for SQL Server 2005 virtual disk implementation because it offers high performance and fault tolerance by combining mirroring and striping.[2] When possible, the disk groups that contain tempdb, the transaction logs, and virtual disks for the data files should be configured with RAID-10.

Because additional drives are required to implement RAID-10, it may not be the preferred choice for certain databases. In these cases, RAID-0 can be used for the tempdb virtual disk. If RAID-0 is used, however, loss of a tempdb disk will affect system availability.

For the transaction log virtual disk, RAID-1 can be used instead of RAID-10 because it provides protection against drive hardware failure. For the virtual disks for the data files, RAID-5 provides a cost-effective alternative, especially for predominantly read-only workloads such as a data warehouse database. However, RAID-5 is not suitable for heavy write workloads such as an online transaction processing database, because it can have significantly lower write performance than other RAID levels due to the reading and writing of parity blocks in addition to the reading and writing of database data.

Configuring the OS and the file system

As mentioned in the preceding section, Microsoft recommends that the number of

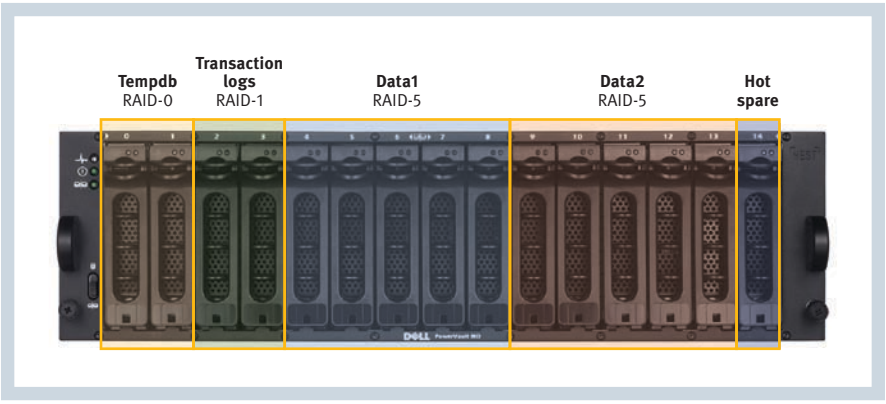


Figure 2. *Separation of tempdb, transaction logs, and data files onto distinct physical disks within a Dell PowerVault MD3000 storage array*

[1] "Storage Top 10 Best Practices," by Microsoft, *Microsoft TechNet*, October 17, 2006, www.microsoft.com/technet/prodtechnol/sql/bestpractice/storage-top-10.mspx.

[2] "Physical Database Storage Design," by Kathy Lu and Lewis Bruck, *Microsoft TechNet*, February 12, 2007, www.microsoft.com/technet/prodtechnol/sql/2005/physdbstor.mspx.

Virtual disk	Minimum size	RAID level	Use
Data1	25% of the database size	RAID-10; for read-only, RAID-5	Data
Data2	25% of the database size	RAID-10; for read-only, RAID-5	Data
Data3	25% of the database size	RAID-10; for read-only, RAID-5	Data
Data4	25% of the database size	RAID-10; for read-only, RAID-5	Data
Tempdb	Variable (depends on the required tempdb size)	RAID-10 or RAID-0	Tempdb
Transaction logs	10% of the database size	RAID-10 or RAID-1	Logs

Figure 3. *Recommended disk configurations for four-socket servers running Microsoft SQL Server 2005*

Drive letter	Folder names	Format	Use
G:	MSS_DATA_1, MSS_DATA_2	NTFS 64 KB	Data
H:	MSS_DATA_3, MSS_DATA_4	NTFS 64 KB	Data
I:	MSS_DATA_5, MSS_DATA_6	NTFS 64 KB	Data
J:	MSS_DATA_7, MSS_DATA_8	NTFS 64 KB	Data
K:	TEMPDB_DATA_1, TEMPDB_DATA_2, TEMPDB_DATA_3, TEMPDB_DATA_4, TEMPDB_DATA_5, TEMPDB_DATA_6, TEMPDB_DATA_7, TEMPDB_DATA_8	NTFS 64 KB	Tempdb
L:	MSS_LOG1, TEMPDB_LOG1	NTFS 64 KB	Logs

Figure 4. *Sample file system configurations for four-socket, dual-core servers running Microsoft SQL Server 2005*

data files configured for a SQL Server 2005 database equal the total number of processor cores installed on the server to achieve optimal performance and scalability for heavy workloads. Also, in the configurations described in this article, the number of virtual disks for data is equal to the number of physical processor sockets. Each virtual disk created in the storage system is mapped to an OS file partition. Therefore, within each partition used for data, the recommended number of data files should equal the number of cores per processor socket. For example, a four-socket, dual-core system should have two data files for each data partition. The number of tempdb data files should also equal the number of processor cores.

Each OS file partition is assigned a unique drive letter. Microsoft best practices recommend that all partitions be formatted using NT File System (NTFS) with 64 KB clusters. Figure 4 illustrates sample file system configurations for four-socket, dual-core host servers.

Using Microsoft Windows Address Windowing Extensions

Microsoft Windows Address Windowing Extensions (AWE) extend Windows memory management to allow applications to address more memory than the standard 32-bit addressing space of 2–3 GB. With AWE, SQL Server 2005 can support up to 64 GB of physical memory on Windows Server 2003.

AWE applies to 32-bit operating systems with more than 4 GB of physical memory. By default, the 32-bit Windows Server 2003 OS reserves 2 GB of memory for the OS and leaves 2 GB of available process address space—known as user-mode virtual address space—to applications. Windows Server 2003 has a boot.ini switch called /3GB that provides applications access to 3 GB of process address space and limits the OS to 1 GB of physical memory. For a system with less than 16 GB of physical memory, a /3GB switch and a Physical Address Extension switch (/PAE) should be added to the boot.ini file. For a system with more than 16 GB of physical memory, a /PAE switch should be added to the boot.ini file.

To enable AWE for SQL Server 2005, administrators should make sure that the "Maximize data throughput for network applications" option is selected in the Network Connections section of the Control Panel. This option may limit the available memory to SQL Server operations because it gives priority to applications that perform buffered I/O operations by caching their I/O pages in system cache.

Configuring Microsoft SQL Server 2005

Installing SQL Server 2005 builds a single tempdb data file by default. When the database is initially created, the tempdb files are placed on the server's internal drives. Microsoft best practices recommend moving the tempdb files to external disk drives dedicated for tempdb files. Also, creating additional tempdb data files can help avoid latch contention on allocation pages and help solve I/O performance

"By combining Dell PowerEdge servers with Dell PowerVault MD3000 storage systems, administrators can easily deploy PowerEdge servers as building blocks of database solutions."

"The best practices described here are intended to help achieve optimal performance of SQL Server 2005, help simplify operations, improve resource utilization, and cost-effectively scale systems as an organization's needs change over time."

issues. The recommended number of tempdb data files should match the number of processor cores of the server. To help effectively utilize the allocation mechanism, all tempdb data files should be equal in size.

By default, the autogrow option is enabled for tempdb files. However, expanding tempdb too frequently can lead to performance degradation. To avoid this issue, Microsoft recommends pre-allocating the tempdb space with enough space to accommodate the expected workload and setting the `FILEGROWTH` increment large enough to minimize tempdb expansions. The Microsoft-recommended `FILEGROWTH` increment setting is 10 percent.[3]

When a new data file is created or a data file expands as a result of the autogrow function, the pages in the file are initialized—zeros are automatically written to the file before it can be used. This file-zeroing process can be time-consuming and can lead to application time-out, especially during the autogrow activity.

One of the major improvements in SQL Server 2005 is the instant file initialization feature. This feature skips the file-zeroing process; instead, the pages in the data file are overwritten with new data when pages are actually allocated. In February 2007, Microsoft conducted in-house tests of file creation and growth processes that showed a significant performance increase when using this feature. For example, the time required to add a 100 MB file dropped from 2.8 seconds to 0.2 seconds with instant data file initialization.[4]

Enabling Address Windowing Extensions

To enable AWE on SQL Server 2005, the SQL Server 2005 engine must run under a Windows account that has the "Lock pages in memory" option assigned (this option is disabled by default). Enabling this option allows SQL Server 2005 to keep data in physical memory and prevents it from sending data to virtual memory. Next, administrators should use the `sp_configure` command to set the `awe enabled` option to 1, and restart SQL Server 2005.

Following best practices for optimal database performance

Maximizing the performance of databases and their supporting storage systems while reducing management complexity can be critical for a growing enterprise. This article provides a blueprint for configuring a standalone Microsoft SQL Server 2005 database on Dell PowerEdge servers and Dell PowerVault MD3000 storage arrays. The best practices described here are intended to help achieve optimal performance of SQL Server 2005, simplify operations, increase resource utilization, and cost-effectively scale systems as an organization's needs change over time.

Wendy Chen *is a systems engineer with the Dell Database, Applications, and High-Availability team. Her fields of interest include Oracle® database backup and recovery, disaster recovery, and Oracle Real Application Clusters. Her experience includes Oracle database administration, application development, and database solutions engineering. Wendy has an M.S. in Computer Science from Brigham Young University.*

Sumner Stewart *is a senior database administrator with Dell IT Global Database Management Systems. He currently consults and supports the Dell manufacturing business segment for global applications. Sumner's primary focuses include multitiered architectures with zero-downtime tolerances and high-performance transaction systems.*

Paul Rubin *is a senior product manager in the Dell Software Solutions marketing group. He is responsible for planning Dell database solutions. Paul has a B.A. in Computer Science from the University of Texas at Austin.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Microsoft SQL Server 2005:
www.microsoft.com/sql

Microsoft SQL Server TechCenter:
technet.microsoft.com/sqlserver

Dell database solutions for Microsoft SQL Server 2005:
DELL.COM/SQL

[3] For more information about configuring tempdb files, see "Working with tempdb in SQL Server 2005," by Wei Xiao, Matt Hink, Mirek, and Sunil Agarwal, *Microsoft TechNet*, June 26, 2006, www.microsoft.com/technet/prodtechnol/sql/2005/workingwithtempdb.mspx.

[4] "Physical Database Storage Design," by Kathy Lu and Lewis Bruck, *Microsoft TechNet*, February 12, 2007, www.microsoft.com/technet/prodtechnol/sql/2005/physdbstor.mspx.

EXPLORING iSCSI AND iSCSI BOOT FOR SAN IMPLEMENTATIONS

BY ROBERT LUSINSKY
DAVID CHRISTENSEN

The ability to boot server operating systems over an Internet SCSI (iSCSI)–based storage area network can offer multiple advantages, including increased system reliability, simplified management, and accelerated restore processes. To help enterprises utilize this functionality, Broadcom and Dell have teamed up to offer enhanced iSCSI boot on Broadcom® Ethernet controllers in Dell™ PowerEdge™ servers.

Related Categories:

Broadcom

Internet SCSI (iSCSI)

Storage

Storage area network (SAN)

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

The Internet SCSI (iSCSI) protocol provides a method of transporting traditional SCSI commands, data, and status messages over standard TCP/IP networks. iSCSI offers a simple, cost-effective, standards-based way for even enterprises with modest budgets to deploy networked storage and gain the advantages of block-level storage consolidation. And because the familiar SCSI interface is maintained over standard networking topologies, iSCSI can help simplify networked storage implementation for both storage and network administrators.

iSCSI boot allows administrators to boot server operating systems over an iSCSI-based storage area network (SAN), helping simplify both server management and the creation, distribution, and maintenance of server images. Now, Broadcom and Dell have teamed up to offer enhanced iSCSI boot functionality on Broadcom Ethernet controllers in Dell PowerEdge servers.

Networked storage in the data center

Using servers on a network to store data is not a new concept, nor is booting a server from a network. Network attached storage (NAS) has existed since the introduction of Network File System (NFS) more than two decades ago. NAS uses file-level protocols that abstract the underlying drive file system to provide networked storage such as shared drives in Linux® or UNIX® operating systems using NFS or in the Microsoft® Windows® OS using Common Internet File System (CIFS). Although NAS is well suited for file sharing, it does not provide block-based functionality, which includes the ability to boot an OS.

Technologies such as Preboot Execution Environment (PXE), Remote Program Load (RPL), and Bootstrap Protocol (BOOTP), which have existed for years, are the current mechanisms for booting directly from a network and are available on Broadcom Ethernet controllers and Dell PowerEdge servers. Although systems can be booted through a combination of NAS and PXE, RPL, and BOOTP, configuring systems to do so is not a simple matter, and may restrict system performance or capabilities.

The introduction of SANs provided block-level storage over a network, allowing shared resources to store application data from multiple servers and thereby helping increase storage utilization while reducing the number of devices for administrators to back up and manage. In this configuration, the file system appears to be present on a local direct attach storage (DAS) disk within the server, but actually exists on a SAN array. SAN technologies such as iSCSI help significantly simplify booting an OS from a

"iSCSI offers a simple, cost-effective, standards-based way for even enterprises with modest budgets to deploy networked storage and gain the advantages of block-level storage consolidation."

network by providing the OS with low-level access to the storage device, which, in turn, allows native file system access and the use of standard OS disk utilities.

Networked storage systems such as NAS and SANs allow organizations to easily divide and share highly available pools of storage among groups of users, helping meet applications' increasing storage needs. These types of storage enable increased utilization, scalability, and simplified backup solutions while helping increase data availability. Multiple enterprise requirements contribute to the need for networked storage:

- **Efficient management:** Consolidating to networked storage allows administrators to manage a single pool of central storage resources, rather than disparate, unconnected storage systems. Consolidation can also provide benefits such as reduced total cost of ownership and easy retargeting of servers, helping create a dynamic, scalable data center.
- **High availability and robust disaster recovery:** Networked storage supports high levels of availability while helping simplify disaster recovery, backup, and data redundancy systems.
- **Regulatory compliance:** Government regulations such as the Sarbanes-Oxley Act and Health Insurance Portability and Accountability Act (HIPAA) have increased enterprise documentation, data security, and auditing requirements. Networked storage can help ease the task of complying with these regulations.

iSCSI SAN architecture

As shown in Figure 1, iSCSI-based SANs comprise two parts: an iSCSI initiator and an iSCSI target. The iSCSI initiator resides on the server or client system and connects to the iSCSI target over an IP network, allowing the physical disk on the target system to appear as a local disk. The iSCSI initiator sends commands to and receives responses from the iSCSI target; conversely, the iSCSI target receives and responds to commands from the iSCSI initiator.

The iSCSI protocol stack consists of a set of standard protocols residing at different Open System Interconnection (OSI) layers: Ethernet as the network interconnect (Layer 2), IP for routing (Layer 3), and TCP for transport (Layer 4), with the iSCSI protocol itself residing at Layer 5 (see Figure 2). iSCSI also allows a SCSI interface to be presented to the OS, simplifying OS and application utilization. In addition, because iSCSI is built on standard TCP/IP protocols, it can be routed over IP version 6 (IPv6) on packet-switched internetworks.

By combining existing and widely deployed standards, iSCSI offers many advantages over other SAN fabrics, including the following:

- **Reduced learning curve:** Standard TCP/IP and Ethernet components are typically familiar to IT staff, helping reduce training, administrative, and maintenance costs.
- **Interoperability:** Utilizing standardized networking technologies helps ensure interoperability between SAN components.
- **Flexibility:** iSCSI supports a wide range of Ethernet network topologies, which can take

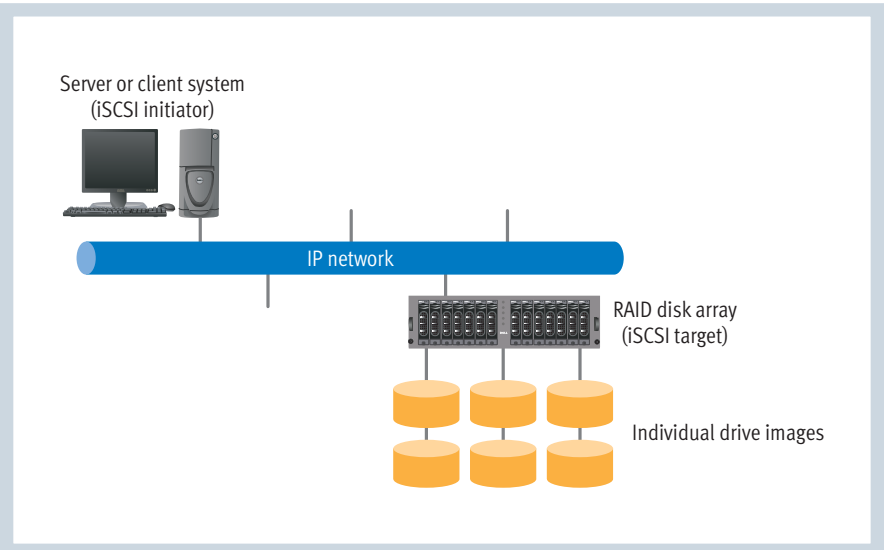


Figure 1. *iSCSI-based SAN architecture*

Protocol	OSI layer	Publication date	Specification
Ethernet	Layer 2	1974	IEEE 802.3
IP	Layer 3	1981	IETF RFC 791
TCP	Layer 4	1981	IETF RFC 793
iSCSI	Layer 5	2004	IETF RFC 3720

Figure 2. *Protocols used for iSCSI boot*

advantage of the routing capabilities inherent in the IP protocol. In addition, as a routable networked storage system, an iSCSI-based SAN can be used over the Internet and across the enterprise.

- **Cost-effectiveness and simplified management:** Using a single fabric for networking and storage helps avoid the additional acquisition and maintenance costs and administrative overhead of managing a separate fabric dedicated to storage.
- **Virtual LAN (VLAN) support:** VLANs help isolate network and storage traffic and support quality-of-service protocols on Ethernet.

These advantages have combined to accelerate the adoption of iSCSI—for example, iSCSI initiators are now a standard part of both Microsoft Windows and Linux operating systems.

iSCSI boot functionality over SANs

iSCSI boot allows a server to boot an OS over a SAN, helping eliminate the need for local disk storage, enhancing system reliability, and simplifying administrator workloads by centralizing the creation, distribution, and maintenance of server images. iSCSI boot can also accelerate server restore processes following a server failure: rather than having to re-image the replacement server, administrators can simply point the new server to the boot image on the SAN to quickly bring it online.

Deploying iSCSI to boot an OS over a SAN can be a relatively simple process. It requires the following system components:

- iSCSI boot–capable Ethernet controller with iSCSI boot ROM supporting the iSCSI, SCSI, IP, and TCP protocols
- iSCSI boot ROM to hook the BIOS Interrupt 13h interface and provide Extended Interrupt 13h interface services complying with Enhanced Disk Drive Specification 3.0, to allow support for drives larger than 2 GB through the use of 64-bit logical block addressing
- BIOS Boot Specification (BBS) support to allow the iSCSI boot implementation to seamlessly integrate with the BIOS boot process and boot selection interface
- OS-based iSCSI initiator, such as the Microsoft iSCSI Software Initiator for Microsoft Windows or the Open-iSCSI Software Initiator for Linux

The actual iSCSI boot process is divided into two phases: real mode and protected mode. These modes refer to the x86 processor operating mode, and roughly equate to the system BIOS and OS, respectively. In real mode, iSCSI boot transactions in systems with Broadcom Ethernet controllers are handled by the Broadcom iSCSI boot initiator. The BIOS uses this initiator to load the disk's master boot record (MBR), the OS bootstrap loader, and the OS bootstrap loader, then uses the Broadcom iSCSI boot initiator to load the OS kernel and the necessary OS drivers. Once the OS kernel is initialized and running, the processor transitions from real mode to protected mode and begins using the OS native drivers, including the OS iSCSI initiator, to complete the iSCSI boot process.

One exception to this boot process occurs when booting the Microsoft MS-DOS® OS using iSCSI boot. Because MS-DOS does not use the processor's protected mode, MS-DOS continues using the Broadcom iSCSI boot initiator for disk transactions even after it has booted.

One of the challenges of the iSCSI boot process relates to passing the iSCSI boot parameters and network settings from the Broadcom iSCSI boot initiator to the OS iSCSI initiator. Proprietary solutions have existed for some time, but today, the industry is starting to adopt the Microsoft iSCSI Boot Firmware Table (iBFT) specification, a table of iSCSI and network parameters written in memory by the iSCSI boot initiator and read by the OS iSCSI

"iSCSI boot allows a server to boot an OS over a SAN, helping eliminate the need for local disk storage, enhancing system reliability, and simplifying administrator workloads by centralizing the creation, distribution, and maintenance of server images."

"Broadcom and Dell have teamed up to create a simple yet richly featured iSCSI boot solution designed to allow iSCSI-based SANs to replace many forms of local storage."

initiator when it initializes. The table serves as a "handoff" between the real-mode and protected-mode phases of the boot process. Microsoft supports iBFT on the Microsoft Windows Server® 2003 OS and the next version of this OS, Windows Server 2008 (code-named "Longhorn"). Open-iSCSI supports iBFT in Linux.

Broadcom iSCSI boot solution on Dell PowerEdge servers

Broadcom and Dell have teamed up to include the Broadcom iSCSI boot solution on Dell PowerEdge servers that use the Broadcom NetXtreme II™ BCM5708 Ethernet controller. This feature is also available on Broadcom NetXtreme II BCM5708–based mezzanine network interface cards for Dell blade servers as well as on Broadcom NetXtreme™ BCM5722 Ethernet controllers.

Key features supported by Dell and Broadcom using iSCSI boot include the following:

- **Single ROM image:** The Broadcom iSCSI boot solution uses a single ROM image to provide iSCSI boot along with PXE, RPL, and BOOTP, helping avoid the need to flash separate ROM images when deploying different network boot technologies.
- **Integrated BBS support:** Booting from an iSCSI disk is as simple as selecting the Broadcom Ethernet controller from the boot menu in the Dell server BIOS.
- **Multiple target support:** Administrators can specify both a primary iSCSI target and a secondary failover target for iSCSI boot, so that if the server cannot establish a connection to the primary iSCSI target, it can use the secondary target instead.
- **Challenge Handshake Authentication Protocol (CHAP) support:** Support for one-way and mutual CHAP authentication helps increase security when establishing a connection between iSCSI initiators and targets.
- **Flexible iSCSI and network parameters:** When configuring iSCSI initiator and target information as well as standard network settings, administrators can either set the parameters statically in the Ethernet controller setup menu or retrieve them dynamically through Dynamic Host Configuration Protocol (DHCP).
- **iSCSI target redirection:** iSCSI targets can redirect the Broadcom iSCSI boot ROM to use a different iSCSI target for the boot process.
- **Linkup time delay:** Administrators can insert a delay during the boot process to accommodate targets that may need additional time to come online after a reset.
- **Microsoft Multipath I/O (MPIO) and device mapper (DM) multipath support:** Broadcom multipathing supports native MPIO in Microsoft Windows and DM multipath in Linux, helping provide network redundancy between iSCSI initiators and targets.
- **Target connections without boot:** Servers can make connections to iSCSI targets without actually booting from the iSCSI boot drive, helping make these targets readily available for direct OS installations.

Simple, flexible networked storage environment

As the use of SANs continues to grow, and the advantages of moving from local storage to centrally managed networked arrays become increasingly important, network boot solutions such as iSCSI boot should become a common feature of enterprise data centers. Broadcom and Dell have teamed up to create a simple yet richly featured iSCSI boot solution designed to allow iSCSI-based SANs to replace many forms of local storage. Deploying this solution can help enterprises of all sizes realize the advantages of iSCSI-based networked storage while helping simplify management and reduce total cost of ownership.

Robert Lusinsky *is a marketing director for the High Speed Controller Business Unit at Broadcom. Robert has a B.S. and an M.S. in Electrical Engineering from California State University and an M.B.A. from the University of Southern California.*

David Christensen *is an applications engineering manager for the High Speed Controller Business Unit at Broadcom. David has a B.S. in Computer Science from California State University, Fullerton, and has worked at a variety of software engineering and product support roles at AST Computer, Phoenix Technologies, and Broadcom over the last 17 years.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Broadcom:
www.broadcom.com

Dell PowerEdge servers:
DELL.COM/PowerEdge

Dell iSCSI storage solutions:
DELL.COM/iSCSI

OPTIMIZED DELL POWERVAULT STORAGE FOR MICROSOFT EXCHANGE SERVER 2007

BY SUMAN KUMAR SINGH
MAHMOUD AHMADIAN

Dell-recommended solutions for Microsoft® Exchange environments can help simplify the design and deployment of messaging infrastructures using industry-standard components. This article describes two optimized solutions using Serial Attached SCSI (SAS)–based Dell™ PowerVault™ MD1000 and PowerVault MD3000 storage.

Related Categories:

Dell PowerVault storage

Microsoft Exchange Server 2007

Serial Attached SCSI (SAS) Storage

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

E-mail applications have become one of the most vital elements of enterprise IT, and the performance and availability of messaging infrastructures can be critical. Choosing the appropriate platform for messaging software such as Microsoft Exchange, however, can be a complex task. To help reduce this complexity for enterprises of all sizes, Dell has performed extensive testing and characterization of Exchange Server 2007 on standard Dell building blocks.

The Exchange Solution Reviewed Program (ESRP) was developed by Microsoft to provide a common storage testing framework that allows vendors to provide information on storage for Exchange. It integrates both the Microsoft Exchange Server Jetstress storage testing tool and solution-publishing guidelines. Dell utilizes the ESRP to test and optimize its storage for Exchange. This article presents two such recommended solutions using Serial Attached SCSI (SAS)–based Dell PowerVault MD1000 and PowerVault MD3000 storage.

Organizations should keep in mind that the ESRP is not a benchmarking program, and its tests are not designed to indicate maximum throughput for a given solution; rather, it focuses on producing recommended solutions for optimized Exchange deployments. The solutions presented in this article are only a small subset of the available ESRP solutions recommended by Dell. A complete list is available at DELL.COM/Exchange.

Dell PowerVault MD1000–based storage solution for Microsoft Exchange

The PowerVault MD1000–based storage solution for Exchange Server 2007 has been tested and validated to support 2,000 users and utilize cluster continuous replication (CCR), a high-availability clustering mechanism based on the Microsoft Cluster Service majority node set model. CCR requires an active and a passive set of Exchange Mailbox server nodes, each of which maintains its own copy of the mailbox databases. E-mail clients access the primary (active) server, and database changes to this server are sent to the secondary (passive) server in the form of log records. The log records are played on the secondary server to help keep the secondary database consistent with the primary database.

The PowerVault MD1000 is a modular disk storage expansion enclosure for Dell PowerEdge™ servers that can house up to fifteen 3.5-inch disk drives in a single 3U rackable chassis. This direct attach storage (DAS) enclosure supports both SAS and Serial ATA (SATA) disk drives. The PowerEdge Expandable RAID Controller (PERC) 5/E connects the PowerVault MD1000 storage enclosure to the server, and supports 3 Gbps SAS as the storage interconnect and PCI Express (PCIe) as the host-based interconnect.

This solution includes a primary site and a secondary site, both with a single PowerEdge 2950 server attached to a PowerVault MD1000 storage enclosure. The primary and secondary storage do not share storage array controllers or

disks. Figures 1 and 2 outline the components and configuration of this solution.

The tested user profile was for 2,000 mailboxes, assuming a 300 MB mailbox size and 0.5 I/Os per second (IOPS) for each user (0.42 IOPS per user, with room for a 20 percent increase). The storage configuration was as follows:

- **Disks 0–9:** RAID-10 volume for Exchange Information Stores
- **Disks 10–11:** RAID-1 volume for Exchange transaction logs
- **Disk 12:** Hot spare
- **Disks 13–14:** Available for future growth

This solution is designed with scalability in mind to help accommodate future growth. Organizations can easily scale it out either by daisy-chaining additional enclosures or by adding server and storage building blocks.

Servers	
Servers	One Dell PowerEdge 2950 server at each site
Processors	Two dual-core Intel® Xeon® 5160 processors at 3.00 GHz per server
Memory	16 GB of double data rate 2 (DDR2) error-correcting code (ECC) memory per server
Internal disks	Two 73 GB, 15,000 rpm Seagate ST373454SS SAS drives per server
Network interface card (NIC)	One Broadcom NetXtreme II NIC per server
RAID controller	One PERC 5/i (firmware version 1.00.01-0088) per server
Storage	
Storage	One Dell PowerVault MD1000 at each site
Disks	Fifteen 300 GB, 15,000 rpm Seagate ST33006555ss/Rev S512 SAS drives per array
RAID controller	One PERC 5/E (firmware version 1.03.10-0216) per array

Figure 1. *Components of validated Dell PowerVault MD1000–based storage solution for Microsoft Exchange*

Performance tests: I/O, backup and recovery, and log reads

In July 2007, Dell engineers conducted performance tests on the PowerVault MD1000–based storage solution for Exchange Server 2007 using the configuration described in the preceding section. Figure 3 summarizes the results. The database and transaction log I/O tests were designed to use the storage with the maximum sustainable level of Exchange I/O, to show how long the storage takes to respond to I/O under load. The data in Figure 3 is the sum of all the logical disk I/Os and the average of all the logical disks' I/O latency during the test.

The streaming backup and recovery test was designed to measure the maximum rate at which databases could be backed up. The test team measured the database and log read I/O performance metrics by running a checksum on the databases and log files. Figure 3 shows the average rate for a single database file.

The log read test was designed to measure the maximum rate at which the log files could be played against the databases. Figure 3 shows the average rate for 500 log files played in a single storage group, where each log file was 1 MB in size.

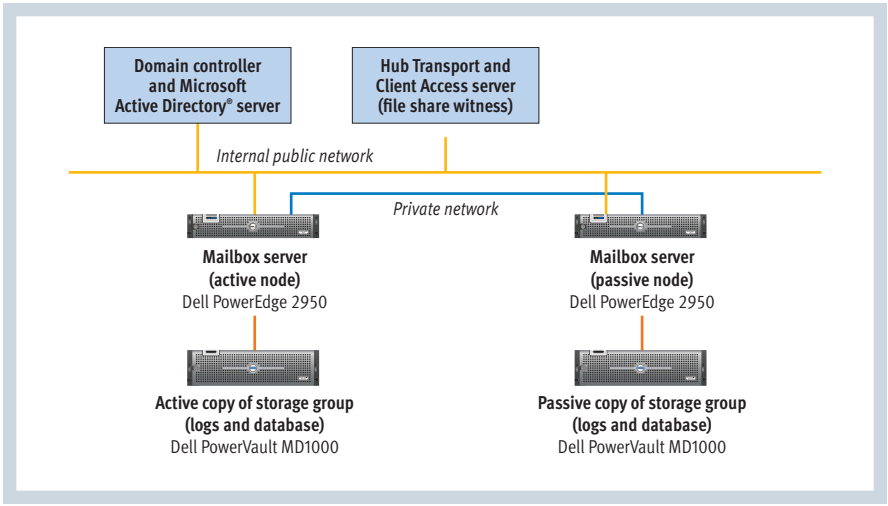


Figure 2. *Configuration of validated Dell PowerVault MD1000–based storage solution for Microsoft Exchange*

Database I/O	
Average database disk transfers/sec	1,435
Average database disk reads/sec	745
Average database disk writes/sec	690
Average database disk read latency	16 ms
Average database disk write latency	4 ms
Transaction log I/O	
Average log disk writes/sec	477
Average log disk write latency	0 ms
Streaming backup and recovery	
Average 1 MB reads/sec per storage group	57
Total 1 MB reads/sec	286
Log reads	
Average time to play one log file	0.36 sec

Figure 3. *Performance results for the validated Dell PowerVault MD1000–based storage solution for Microsoft Exchange*

Servers	
Servers	Two Dell PowerEdge 2950 servers
Processors	Two dual-core Intel Xeon 5160 processors at 3.00 GHz per server
Memory	16 GB of DDR2 ECC memory per server
Internal disks	Two 73 GB, 15,000 rpm Seagate ST373454SS SAS drives per server
NIC	One Broadcom NetXtreme II NIC per server
RAID controller	One PERC 5/i (firmware version 1.00.01-0088) per server
Storage	
Storage	Dell PowerVault MD3000 (firmware version 06.17.77.60), with two PowerVault MD1000 expansion enclosures
Disks	Fifteen 300 GB, 15,000 rpm Seagate ST33006555ss/Rev S512 SAS drives per array
RAID Controller	One PERC 5/E (firmware version 00.10.49.00) for the PowerVault MD3000

Figure 4. *Components of validated Dell PowerVault MD3000–based storage solution for Microsoft Exchange*

Dell PowerVault MD3000–based storage solution for Microsoft Exchange

The PowerVault MD3000–based storage solution for Exchange Server 2007 has been tested and validated to support 3,000 users and utilize single copy clustering (SCC), a high-availability clustering mechanism based on Microsoft Cluster Service shared storage. SCC requires an active and a passive set of Exchange Mailbox server nodes sharing the mailbox databases. During normal operations, e-mail clients access the primary (active) server. If the primary node fails, the secondary (passive) node takes over as the active node. Compared with previous versions of Exchange software, Exchange Server 2007 has enhanced deployment setup and management functionality for this clustering model.

The PowerVault MD3000 is a modular disk storage system that can expand with up to two PowerVault MD1000 storage enclosures. The PERC 5/E connects the PowerVault MD3000 to the server, and supports 3 Gbps SAS as the storage interconnect and PCIe as the host-based interconnect.

This solution includes two PowerEdge 2950 servers attached to a PowerVault MD3000 storage system. The primary (active) and secondary (passive) nodes use the same configuration, with the primary node providing client access to the Exchange Information Store. The storage is shared between these two nodes. Figures 4 and 5 outline the components and configuration of this solution.

The tested user profile was for 3,000 mailboxes, assuming a 1 GB mailbox size and 0.5 IOPS per user (0.42 IOPS per user, with room for a 20 percent increase). The storage configuration in each of the three storage enclosures was as follows:

- **Disks 0–9:** RAID-5 volume for Exchange Information Stores
- **Disks 10–11:** RAID-1 volume for Exchange transaction logs
- **Disk 12:** Hot spare
- **Disks 13–14:** Available for future growth

Like the PowerVault MD1000–based solution, organizations can scale this solution out by adding server and storage building blocks.

Performance tests: I/O, backup and recovery, log reads, and RAID level

In August 2007, Dell engineers conducted performance tests on the PowerVault MD3000–based storage solution for Exchange Server 2007 using the configuration described in the preceding section. Figure 6 summarizes the results of these tests, which had the same design and goals as those performed on the PowerVault MD1000–based solution.

The ESRP focuses on storage testing to help address performance and reliability issues with storage design; however, storage is not the only factor to take into consideration when designing a scalable Exchange solution. For example, the PowerVault MD3000–based storage solution for Exchange uses RAID-5 containers for the Exchange database, which can be effective for environments with very large mailboxes and/or

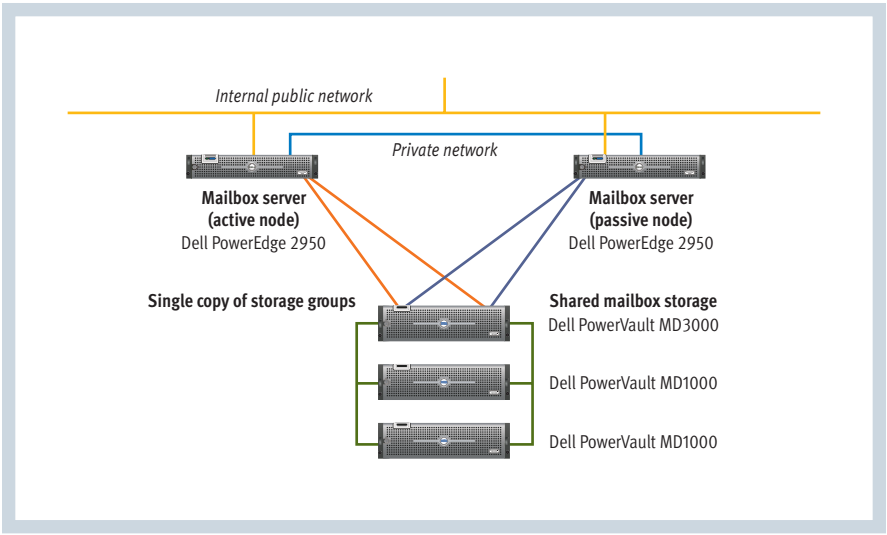


Figure 5. *Configuration of validated Dell PowerVault MD3000–based storage solution for Microsoft Exchange*

"E-mail applications have become one of the most vital elements of enterprise IT, and the performance and availability of messaging infrastructures can be critical."

low IOPS requirements. RAID-10 typically offers the highest performance with high levels of protection, but only leaves half of the RAID group's capacity usable. RAID-5 offers a higher usable capacity per RAID group than RAID-10, but lower I/O throughput.

There are two reasons for this reduced throughput. First, under normal operating conditions, RAID-5 requires the extra work of generating and updating parity information, which is what makes RAID-5 fault tolerant and RAID recovery possible. Second, following a disk failure, RAID-5 containers must restore data and calculate its parity in addition to serving normal I/O requests, which can significantly reduce the performance during rebuild operations. In addition, a RAID-5 volume can only tolerate a single disk failure, while RAID-10 can survive up to *n* disk failures for an *n* + *n* RAID-10 volume.

As shown in Figure 7, performance can suffer significantly during a RAID-5 rebuild cycle. In an optimal state—before failure—the RAID-5 database disks provided 1,709 IOPS with a read latency of 11 milliseconds and a write latency of 9 milliseconds. In a degraded state, the I/O throughput was reduced to 1,545 IOPS and the database read latency increased to 14 milliseconds. During the rebuild cycle, the throughput was further reduced to 1,116 IOPS—approximately 35 percent less than the throughput in the optimal state—with a read latency of 18 milliseconds and a write latency of 16 milliseconds.

A variation of the PowerVault MD3000–based solution would use RAID-10 for the Exchange databases, which could provide higher throughput than RAID-5. However, this approach would also reduce the total size of the database logical units from 7,344 GB to approximately 4,080 GB, thereby limiting the user mailbox size to 650 MB rather than 1 GB.

Optimized storage for Microsoft Exchange

SAS-based Dell PowerVault systems provide a standards-based storage platform for Microsoft Exchange deployments. By taking advantage of the Dell-recommended PowerVault MD1000– and PowerVault MD3000–based solutions described in this article, enterprises can help simplify the design and deployment of their Exchange Server 2007 infrastructure while helping optimize messaging performance.

Suman Kumar Singh *is a lead engineer on the Dell End-to-End Solutions team. He specializes in messaging systems architecture and sizing, and led the release of the Dell Exchange 2003 Advisor tool. His other interests include storage area networks (SANs), virtualization, and security, and he has published and presented several papers at industry conferences. Suman has a master's degree from the University of Texas at El Paso.*

Mahmoud Ahmadian *is an engineering consultant on the Dell End-to-End Solutions team. His technical interests include clustering technologies, databases, and messaging systems performance. He has an M.S. in Computer Science from the University of Houston, Clear Lake.*

	Optimal	Degraded	Rebuilding
Database IOPS	1,709	1,545	1,116
Read latency	11 ms	14 ms	18 ms
Write latency	9 ms	9 ms	16 ms

Figure 7. *Test results for RAID-5 performance in the validated Dell PowerVault MD3000–based storage solution for Microsoft Exchange*

Database I/O	
Average database disk transfers/sec	1,709
Average database disk reads/sec	887
Average database disk writes/sec	822
Average database disk read latency	12 ms
Average database disk write latency	11 ms
Transaction log I/O	
Average log disk writes/sec	518
Average log disk write latency	2 ms
Streaming backup and recovery	
Average 1 MB reads/sec per storage group	22.4
Total 1 MB reads/sec	336
Log reads	
Average time to play one log file	0.50 sec

Figure 6. *Performance results for the validated Dell PowerVault MD3000–based storage solution for Microsoft Exchange*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINK

Microsoft Exchange Server 2007 solutions on Dell:
DELL.COM/Exchange

ENTERPRISE-CLASS SAN SOLUTIONS FOR SMALL AND MEDIUM BUSINESSES

BY BALA NAGESH

Storage area networks (SANs) have traditionally been reserved for large enterprises with the budgets and IT staff resources to handle them. Now, Dell and Cisco have collaborated to create SAN solutions designed specifically for small and medium businesses, allowing them to deploy high-performance, enterprise-class storage in a cost-effective, easy-to-manage way.

Related Categories:

Cisco Systems

Storage

Storage area network (SAN)

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

The collaboration between Dell and Cisco stems in part from their shared commitment to small and medium businesses (SMBs). SMBs often need the advantages of enterprise-class systems, but face data management challenges magnified by faster growth rates, smaller budgets, and smaller IT staffs than are typically present in large enterprises.

Although SMBs can clearly benefit from consolidating resources in storage area networks (SANs), many obstacles make it difficult for them to do so. For example, the complexity and high cost of entry of SANs can be a major roadblock, and SMBs that implement them can face similarly high ongoing costs for support and maintenance.

The Cisco® MDS 9124 multilayer fabric switch can help SMBs meet these challenges and enable a painless transition to a SAN (see Figure 1). To help simplify adoption even further for SMBs, Dell is working with Cisco to offer pretested, preconfigured SAN solutions built on Dell™ hardware and the Cisco MDS 9124 switch.

Creating integrated SAN solutions for small and medium businesses

In the past, SMBs have typically had to choose between low-cost, entry-level SAN systems that limited expandability, and expensive, high-capacity SAN systems that required purchasing more hardware than necessary. Dell SMB SAN solutions help eliminate this trade-off, enabling a SAN strategy that can scale appropriately as needed.

Combining Dell hardware with the Cisco MDS 9124 switch, these solutions incorporate simplified selection, installation, and management to help SMBs deploy enterprise-class storage without a dedicated team of storage experts. And the building-block approach of these solutions lets SMBs easily configure server and SAN components to meet capacity and feature requirements.

The Cisco MDS 9124 SAN switch, a key component of these solutions, has been designed with SMBs in mind. Its features include the following:

- **High performance:** The switch can deliver throughput of up to 4 Gbps on each port and up to 192 Gbps aggregate with 24-port configurations. This level of performance enables the switch to work with both current and future high-performance drives, helping maximize its useful lifetime and increase its long-term value.
- **High port density:** With 8–16 ports or 16–24 ports in each 1U rack-mountable switch, Dell SMB SAN solutions are designed to fit into small offices and minimize power and cooling requirements (see Figure 2). The switch's density/price ratio provides cost-effectiveness while allowing administrators to add to the SAN as needed.

- **High availability:** A reduced number of internal parts helps increase reliability, while redundant hot-swappable power supplies contribute to high availability. Administrators can also integrate multiple switches into each SAN to avoid single points of failure. To help avoid fabric disruptions, the switch supports adding servers and storage through on-demand port activation as well as performing common configuration tasks—such as upgrading firmware or optical transceivers—without taking the switch offline.
- **Multi-layer fabric support:** The switch supports both Fibre Channel and virtual SANs (VSANs). VSANs enable administrators to partition a switch into multiple logical or virtual switches, and each VSAN can have its own fabric services and zoning.
- **Simplified setup and management:** An integrated configuration wizard that lists servers and storage allows administrators to quickly bring SANs online. Administrators can also use the point-and-click provisioning menu to easily configure zoning, VSANs, and other configuration options without having to know device World Wide Names (WWNs). In addition, the pretested configurations of Dell SMB SAN solutions incorporating the Cisco MDS 9124 switch help greatly simplify deployment, and integrated graphical SAN management tools and diagnostics help reduce the time required for ongoing management.

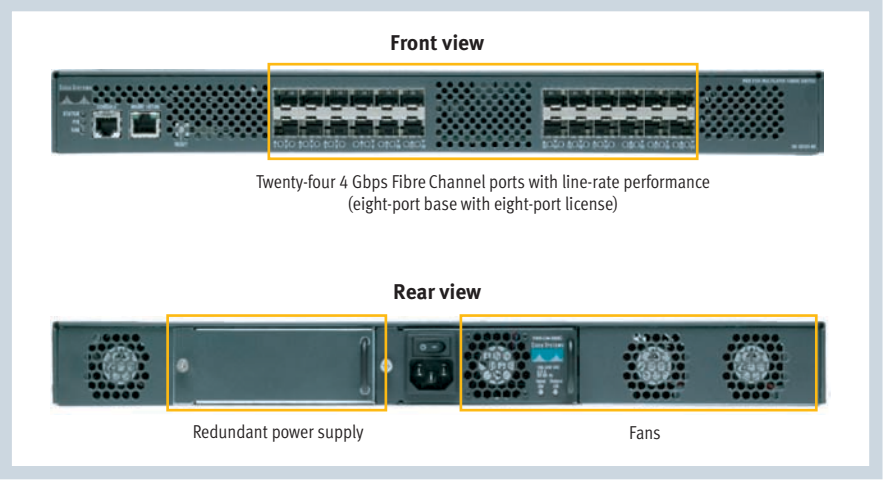


Figure 2. *Front and rear views of the Cisco MDS 9124 switch*

- **Integrated security:** The switch includes features such as secure management access control, role-based access control, hardware-enforced zoning, physical port lockdown, and audit trails to monitor configuration and policy changes.

Enabling cost-effective, enterprise-class SANs

Dell SMB SAN solutions with the Cisco MDS 9124 switch are designed to help SMBs deploy and scale enterprise-class storage cost-effectively while minimizing the management complexity and cost overhead that can be impediments for small IT teams. In addition, because the collaboration between Dell and Cisco has created an end-to-end supply chain, SMBs can order SAN solutions through familiar channels, with the solutions then integrated at the factory and preconfigured for rapid deployment. These solutions are complemented by world-class support from Dell service teams backed by Cisco global support organizations. Dell and Cisco both hope that their collaboration will continue to provide industry-leading joint solutions like these in the future for both SMBs and other enterprises.

Bala Nagesh *is a product line manager for MDS fabric switches at Cisco Systems.*

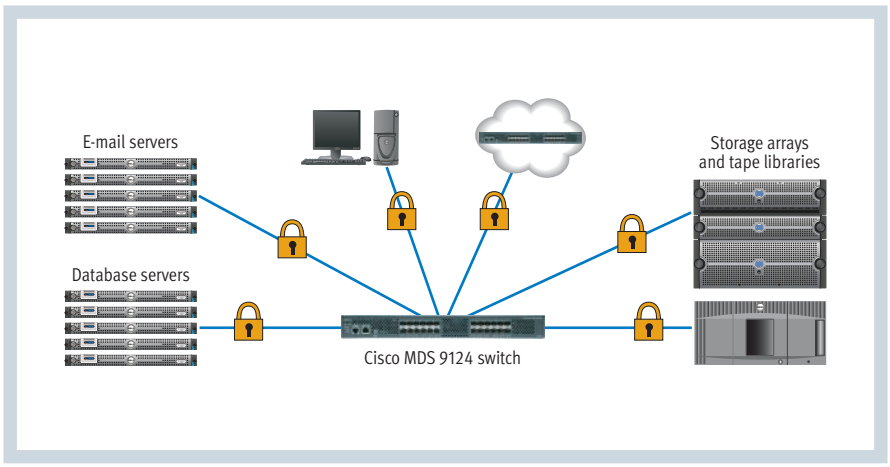


Figure 1. *Cisco MDS 9124 multilayer fabric switch in an SMB environment*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell SMB solutions:
DELL.COM/SMB

Cisco MDS 9124 switch for Dell SMB SAN solutions:
DELL.COM/content/products/productdetails.aspx/cisco_mds9124

Cisco storage switches:
www.cisco.com/go/storage

PROTECTING VITAL BUSINESS INTERESTS WITH A SIMPLIFIED STORAGE SOLUTION

A flood tide of digital data and increasingly complex regulatory requirements have made data backup, recovery, and archiving critical aspects of enterprise operations. Dell Services experts can assess how effectively an information infrastructure is getting the job done and deploy integrated, best-of-breed storage solutions to help ensure the safety and recoverability of critical data.

In today's enterprise environment, the amount of digital data being created and stored is growing at an astonishing rate. Enterprises are not only relying on IT to enhance the quality and efficiency of key functions such as customer support and manufacturing; they are also being compelled by regulatory requirements to retain vast amounts of information.

To help maintain business continuity and mitigate the risk of disaster or litigation, enterprises must have the agility to retain, protect, and recover ever-increasing volumes of data quickly, flexibly, and cost-effectively. It is no longer enough to have an adequate backup and recovery process in place to access data and restore enterprise operations in the event of a disaster, outage, or accidental loss. Today, organizations also need an archiving strategy designed to ensure that they can retain and manage data over the long term to satisfy regulatory, auditing, litigation, and other data management requirements.

Demystifying complex storage requirements

To help enterprises assess how well their information infrastructures meet their data backup, recovery, and archiving needs, and to help them deploy cost-effective, best-of-breed solutions, Dell Services offers a comprehensive suite of services designed to increase data security, reduce backup and recovery times, facilitate regulatory compliance, simplify management, and lower the total cost of backup, recovery, and archiving.

Although the ability to quickly restore enterprise systems following a failure is a critical aspect of day-to-day IT operations, backup and recovery can be complex, inefficient, and costly. As a result, many organizations have not implemented processes for backup and recovery, and those that have often perform these vital functions in an inconsistent and ad hoc fashion.

For example, it is often difficult to determine what data should be backed up, and when. Some data may be mission critical; other data may be relatively unimportant. Some data may be constantly changing; other data may be relatively static. Organizations that do not have a clear idea of their data backup requirements often back up too much data, which can dramatically slow down backup and recovery times and lead to an unnecessary investment in storage resources. Or they may inadvertently neglect to back up critical data, which can put them at serious risk in the event of a failure or audit.

All too often, enterprises that do implement backup and recovery systems end up with a mix of complex and costly storage and network technologies that require a wide range of specialized skills to operate. For example, many backup and recovery systems include legacy disk and tape drives, storage area networks (SANs) and Fibre Channel networking, and other technologies such as network attached storage (NAS) and Ethernet networking.

In addition to implementing recoverability systems and processes, enterprises also must ensure that they

comply with regulatory requirements. These requirements, included in legislation such as the Sarbanes-Oxley Act and the Health Insurance Portability and Accountability Act (HIPAA), demand that companies retain, and be able to produce on demand, an increasing amount of data encompassing e-mail as well as financial records, personnel records, and customer transaction data. When organizations fail to comply with these regulations, they put themselves at serious risk of litigation, fines, and even jail time for executives. One key reason for failure to comply is that regulations are often complicated and hard to interpret. Small and medium-size enterprises typically lack the compliance officers or trained legal staff to help them understand regulatory obligations. Also, like backup and recovery, an inefficient archiving system can be complex and costly.

Assessing specific enterprise needs

Dell Services offers a comprehensive suite to help organizations identify specific storage requirements, assess the efficacy of systems currently in place, and determine whether they should implement backup, recovery, and archiving solutions that are more comprehensive than their current systems.

To begin with, a Dell Services team conducts a half-day workshop to assess what processes and technology are currently in place and zero in on specific challenges the enterprise may need to address. Dell Services then discusses available technology solutions and how they might mitigate areas of concern.

After this initiation phase, a Dell Services team works closely with the IT organization over a span of two to four weeks to gain an in-depth understanding of current backup, recovery, and archiving procedures. As shown in Figure 1, the assessment service includes a data collection phase, a data analysis phase, and a presentation phase. In the data collection phase, a Dell Services team combines a tools- and interview-based approach to assess the organization's current backup, recovery, and archiving procedures. In particular, this team uses software agents to track the overall flow of data within the IT organization. The team also conducts extensive interviews with IT staff to understand both the business and legal requirements for backup, recovery, and archiving and the processes and infrastructure currently in place.

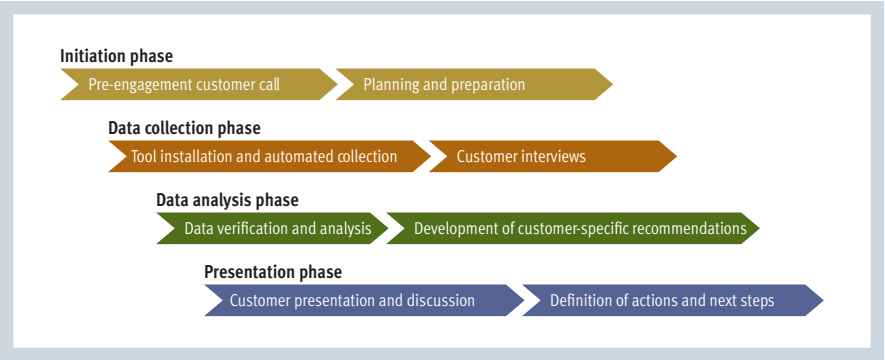


Figure 1. *Dell Backup and Recovery Services assessment methodology*

In the data analysis phase, the Dell Services team verifies and analyzes the data collected during the data collection phase, with the goal of developing specific recommendations. In particular, the team studies the data to determine how well the organization's specific business and legal requirements are being met, as well as to identify any inefficiencies, such as unnecessary backups of duplicate or static data. The team then assesses the overall performance of the current infrastructure, including time to backup, time to recovery, and time to retrieve specific data from archives.

In the presentation phase, the Dell Services team presents specific technology and process recommendations designed to increase data security, reduce backup and recovery times, facilitate regulatory compliance, simplify management, and lower the total cost of backup, recovery, and archiving. This presentation includes an analysis of the total cost of ownership of possible solutions.

For enterprises that decide to implement a backup and recovery and/or archiving solution, Dell also offers a comprehensive design and implementation service. A Dell Services team uses the business requirements and technology and process recommendations from the assessment methodology to design and implement an organization-specific backup, recovery, and archiving solution. The goal of the design and implementation phase is to combine best-of-breed products to create a cost-effective solution that meets bottom-line business requirements.

Enhancing storage performance while avoiding litigation risks

Effective storage systems have become a critical element of enterprise operations. Through its workshop, assessment, design, and implementation services, Dell Services enables organizations to maximize business continuity and mitigate the risk of disaster or litigation by helping them increase the security of essential data, reduce backup and recovery times, facilitate regulatory compliance, simplify management, and lower the total cost of backup, recovery, and archiving.

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell Infrastructure Consulting Services:
DELL.COM/ICS

Dell Backup and Recovery Services:
DELL.COM/Services/Backup&Recovery

SIMPLIFY IT: THE DELL PATH TO INNOVATION AND GROWTH

BY JOE POLLOCK

ENTERPRISES OF ALL SIZES STRUGGLE WITH THE EVER-INCREASING COMPLEXITY OF IT. BY HELPING SIMPLIFY AND STREAMLINE THE REQUIREMENTS FOR ONGOING OPERATIONS AND SYSTEMS MAINTENANCE, THE DELL IT SIMPLIFICATION APPROACH ENABLES ORGANIZATIONS TO RECLAIM THE TIME, MONEY, AND PERSONNEL NEEDED TO TURN INNOVATION INTO A DAILY BUSINESS PRACTICE.

Every day, enterprises around the world struggle with the complexity of IT—and that complexity is increasing. Today, more than ever before, IT professionals have a wide range of responsibilities: aside from maintaining diverse infrastructures that age further each day, they are under constant pressure to rapidly increase IT capabilities to align system functionality with business objectives.

IT professionals are also facing a digital data explosion, with Internet and enterprise users worldwide generating billions of gigabytes of data each year. YouTube, for example, hosts 100 million video streams a day.[1] To exacerbate the situation, smartphones, PDAs, notebooks, tablets, digital cameras, radio frequency identification (RFID) systems, and other devices are blurring the line between business and personal use, further adding to IT staff responsibilities.

Adding to this burden are compliance regulations—such as the Sarbanes-Oxley Act and Health Insurance Portability and Accountability Act (HIPAA) in the United States and the Data Protection Act in the United Kingdom—resulting in many IT organizations storing data for long periods of time. And on top of all this, IT staff must also work constantly to mitigate risks to their operations, ensuring that processes and systems are in place to recover from disasters and protect key assets from security threats.

When trying to meet these varied challenges, the sheer complexity of IT can quickly become unmanageable, regardless of enterprise size or industry. IT simplification then becomes not a luxury, but a critical element of a successful data center—one that can actually enable innovation and spur business growth.

IT COMPLEXITY INHIBITS GROWTH

The more time administrators spend managing and maintaining their IT environment, the less time they can spend developing and implementing new initiatives. In a Forrester report, a survey of 337 IT executives at North American enterprises showed that on average, these enterprises spend about 70 percent of their IT budget on ongoing operations and maintenance, leaving only 30 percent for new initiatives. In that same report, a survey of 452 decision makers at North American enterprises showed that 52 percent of the respondents felt that "driving innovative new market offerings or business practices" was not well supported in their organizations—in fact, this goal was ranked lowest of those listed in the survey. Finally, in a survey of 825 decision makers at North American enterprises, nearly three-quarters cited "improving IT efficiency" as a critical or high business priority for their IT

[1] "YouTube vs. Boob Tube," by Bob Garfield, in *Wired,* December 2006.

> *"The more time administrators spend managing and maintaining their IT environment, the less time they can spend developing and implementing new initiatives."*

organization—by far the highest-ranked option among those listed.[2]

What is the source of IT inefficiency, of these ongoing, day-to-day costs? Time-consuming processes throughout the IT life cycle—from planning to purchase, from installation to implementation, and beyond. A heterogeneous, difficult-to-manage mixture of systems, software, data, interfaces, and applications from a variety of sources, requiring too much manual interaction. Aging, inflexible infrastructures that do not scale well and are too inflexible to meet changing requirements. The constant struggle for IT managers and administrators to keep systems running, patched, and up-to-date; make information accessible when and where it is needed; adhere to complicated licensing and service upgrade requirements with hidden costs; control power and cooling expenses; and keep up with the exponential growth of devices and data.

IT simplification can help reverse this trend. IT managers and administrators should be able to acquire, deploy, manage, and maintain their systems easily, and their IT infrastructures should scale flexibly to meet their needs. As Figure 1 illustrates, reducing ongoing operations and maintenance costs enables enterprises to free up resources, reinvest savings back into the core of the business, build a flexible infrastructure to help meet changing requirements, or reinvent the enterprise altogether.

IT SIMPLIFICATION ENABLES INNOVATION

Over a decade ago, Michael Dell envisioned a data center that ran on standardized systems, and while others touted lofty architectures built on a closed infrastructure, Dell stood alone. The Dell model revolutionized the IT industry, and today, countless enterprises run simple, manageable, scalable IT infrastructures based on x86 standards. This approach is still relevant today.

In fact, Dell itself is built on IT simplification. In 2006 alone, Dell custom-integrated 5.5 million units in its global factories, and it maintains more than 50,000 prepackaged applications as part of the Dell Custom Factory Integration service. The Dell infrastructure includes approximately 130,000 systems worldwide, including 20,000 servers. Despite the global scale of its operations, the company has consolidated 30 data centers down to just 2, and has standardized on only four corporate images: three server images (based on the Microsoft® Windows®, Linux®, and Novell® NetWare® platforms) and one client image. To control this large an infrastructure, Dell manages its servers as aggressively as its desktops—patching and rebooting all Windows-based servers every month, with a success rate of over 95 percent within the first 72 hours for all 130,000 systems. And it relies on only about a dozen people performing these systems management tasks, an assignment that might ordinarily require five times that many.

The efficiency and flexibility of the Dell global infrastructure makes it well suited to help other enterprises simplify their own IT infrastructures. A common set of principles guides Dell in this goal:

- IT should not be as complex as it is today.
- Enterprises should be able to spend less on maintaining IT and more on enabling innovation than they do today. True innovation should become a daily business practice, rather than an ill-defined goal.
- Not every IT project should require an army of consultants. Technology vendors should share knowledge and tools with their customers to enable them to control their own IT.
- IT should have minimal environmental impact.

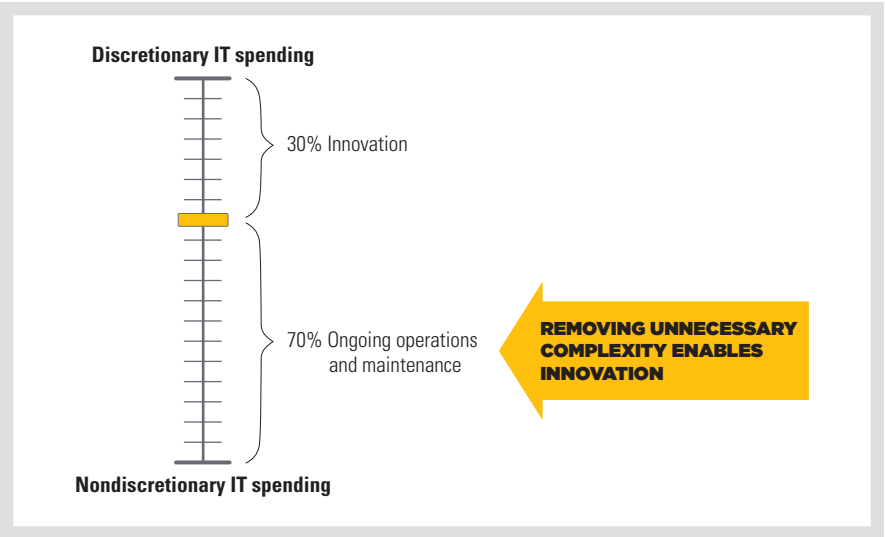


Figure 1. *IT complexity inhibits growth*

[2] "2007 Enterprise IT Budget Outlook: North America," by Forrester, February 8, 2007.

- Enterprises should have superior knowledge about their IT environment, their business, and their needs, enabling them to simplify IT across every aspect of their operations.

Enterprises looking to simplify their own IT should be aware that the high margins, lock-in, and excessive maintenance and support costs that were commonplace decades ago, in the world of RISC-based processors and proprietary UNIX® technology, can still be hidden behind a veil of standards-based offerings from other technology companies. Dell does not lock in customers with proprietary technology and outsourcing contracts, nor does Dell prevent enterprises from unbundling software as they see fit, send out armies of consultants, and ultimately own much of the IT infrastructure.

What Dell did to simplify how PCs were acquired, maintained, and upgraded, it is now applying to all levels of IT—taking the same focus on efficiency that helped Dell build one of the world's most streamlined business models and bringing it to bear on the goal of radically simplifying the practice of IT. The goal is to make IT predictable, reliable, and repeatable, enabling it to run as efficiently as a factory. By combining standardized reference architectures with prevalidated solution stacks, best practices, and specialized services, Dell is working to simplify IT from the desktop to the data center.

CUSTOMIZED ROAD MAPS HELP ENTERPRISES MEET THEIR GOALS

The path to a simplified IT infrastructure must take into account how an organization sees IT. Some view it primarily as an operating cost, using it only for their most essential operations and managing it manually. Some view it as a business enabler, automating it to align with the ever-changing dynamics of their market. And finally, some consider it to *be* their business—a strategic asset that defines their competitive advantage and differentiates them in their market.

Of course, most enterprises think of IT in some combination of these roles, and even people within the same organization might view IT differently depending on their responsibilities. This view frames the meaning of IT simplification. Those who view it as an operating cost might want to acquire as much IT as possible for as little money as possible. Those who view it as a business enabler might want simple tools that can enhance their ability to serve their customers. And those who view it as a strategic asset might seek the latest available technology with the fastest deployment speed.

The Dell approach to IT simplification helps make innovation a daily practice based on an enterprise's needs and how it views IT, with the goal of creating measurable business value. Based on a disciplined, pragmatic approach, IT simplification can streamline deployment, enhance management and maintenance, and provide a foundation that scales easily and cost-effectively to help meet future needs.

The goal of simplicity informs the Dell approach throughout, from products and services to integrated solutions—first by making these elements easy to use; then by making them easy to buy, manage, control, and scale; and finally by basing them on standardized components and platforms to help ensure they can continue to work in the future. Dell looks at the entire technology stack holistically to create an integrated environment that is easy to deploy and manage, and even includes many premium hardware management features at no additional cost. And Dell makes it a goal to share everything it learns with those relying on its products and services, to give those enterprises control over their own technology. Through this approach, Dell enables enterprises to acquire IT quickly, use IT effectively, and scale IT efficiently.

To help understand how IT fits into the needs of the enterprise, and where to focus time and resources, the Dell approach starts with the Dell Simplification Index. This index is designed to objectively measure three key aspects of IT infrastructure:

- **Efficiency:** Efficiency at every stage of the IT life cycle helps enterprises streamline acquisition, deployment, management, maintenance, and growth.
- **Manageability:** The manageability of each layer of the IT stack contributes to IT staff productivity.
- **Flexibility:** IT flexibility helps determine how responsive the environment is to changing needs.

In addition, Dell is currently developing and testing an IT Simplification Assessment Service with enterprises across multiple industries. This service is designed to outline a customized plan with specific steps to help simplify IT over time, asking key questions to identify where an organization's IT stands today, where its IT should be, what the roadblocks are to reaching that goal, and where IT and the enterprise can gain the greatest advantage. For enterprises that view IT merely as an operating cost, the end result might be a plan to strip out cost and complexity to help reduce overall expenses. For those that view IT as a business enabler, it might be to increase efficiency to help the enterprise move forward quickly. For those that view IT as a strategic asset, it might be to significantly reduce the amount of overall spending required for ongoing operations and maintenance, helping increase the resources available for innovation.

After using its IT Simplification Index and IT Simplification Assessment Service to evaluate the infrastructure, identify the areas of greatest

"Dell looks at the entire technology stack holistically to create an integrated environment that is easy to deploy and manage, and even includes many premium hardware management features at no additional cost."

"The Dell approach to IT simplification helps make innovation a daily practice based on an enterprise's needs and how it views IT, with the goal of creating measurable business value."

complexity, and determine a short, cost-effective path to increase efficiency, Dell creates an IT Simplification Plan that is tailored for each customer and includes a set of milestones. These milestones are based on established and agreed-upon goals and take into account both budget and time constraints. This plan is designed to provide benchmarks for success, help organizations make sound decisions, and help them grow in their own way to achieve a high return on investment.

DELL EFFICIENCY BRINGS BOTTOM-LINE ADVANTAGES

Independent research commissioned by Dell has demonstrated the bottom-line advantages of the Dell simplification philosophy. For example, an IDC survey of 200 large enterprises found that deployment activities cost US$527 per PC on average, while costing US$700 per PC or more in specific instances. Companies that optimized their deployments by following the best practices at the core of the Dell PC Deployment Optimization Model, in contrast, saved as much 62 percent (more than US$400) per PC.[3]

Dell™ PowerEdge™ servers can offer similar advantages, not only in deployment time but also in operating efficiency and cost-effectiveness. Studies by Principled Technologies that compared the Dell PowerEdge 1955 blade server system with the HP BladeSystem c-Class blade server system in key categories found that while the Dell system arrived in 2 boxes and took about an hour to deploy, the HP system arrived in 78 boxes and took over three and a half hours to deploy[4]—a significant difference, particularly when scaled to the enterprise level. In a separate evaluation, the Dell PowerEdge 1955 also achieved higher performance per watt than the HP BladeSystem c-Class in every configuration tested.[5] In addition, Dell tests comparing virtualization performance on two-socket Dell PowerEdge 2950 servers and four-socket HP ProLiant DL585 G2 servers found that three Dell PowerEdge 2950 servers provided an average of 44 percent higher performance, 58 percent higher performance per watt, and 95 percent better price/performance than two HP ProLiant DL585 G2 servers.[6]

Enterprises in many different industries have taken advantage of standards-based Dell hardware and services to help simplify the way they acquire, use, and scale their IT infrastructures. For example, using Dell OpenManage™ software, a VMware® virtualization platform, and Dell Services, Welch's was able to migrate to an Oracle® enterprise resource planning solution in less than a week, and reduce provisioning time from three weeks to just 15 minutes using a centralized management console.[7] And companies as varied as Google, Mazda North American Operations, and Unilever have used Dell hardware and services to help simplify and streamline operations while increasing performance and productivity.[8]

DELL IT SIMPLIFICATION HELPS RESHAPE ENTERPRISE IT

The complexity of IT and the sheer volume of data has become a significant challenge, draining resources that could be spent on new initiatives into mundane operations and maintenance tasks. By applying the same focus on efficiency and standards that its own business is built on, Dell can help enterprises of all sizes simplify the acquisition, deployment, management, and maintenance of their IT infrastructures. And once they have reduced the time, money, and personnel required just to keep their systems running, these enterprises are free to reclaim those resources to help increase innovation and core business value.

Joe Pollock *has been with Dell for more than seven years, holding a variety of product management and marketing roles in the Dell Global Product Group and Global Commercial Marketing Organization. He has a B.S. in Electrical Engineering from the University of Florida.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINK

Dell IT simplification:
DELL.COM/Simplify

[3] IDC white paper sponsored by Dell, "Dell PC Deployment Optimization Model," Doc #205282, January 2007.

[4] "Out-of-Box Comparison Between Dell and HP Blade Servers," by Principled Technologies, June 2007, DELL.COM/downloads/global/products/pedge/en/DellHPBladeServerOOB.pdf.

[5] "SPECjbb2005 Performance and Power Consumption on Dell and HP Blade Servers," by Principled Technologies, June 2007, DELL.COM/downloads/global/products/pedge/en/DellHPBladeServerPerfPwr_report.pdf.

[6] "Comparing Virtualization Performance: Dell PowerEdge and HP ProLiant Servers," by Todd Muirhead; Dave Jaffe, Ph.D.; and Terry Schroeder, in *Dell Power Solutions*, May 2007, DELL.COM/downloads/global/power/ps2q07-20070339-Muirhead.pdf.

[7] "Welch's Juices Up Its IT Infrastructure," by Dell Inc., January 2006, DELL.COM/downloads/global/casestudies/461_2005_Welchs.pdf. Results not typical.

[8] For more information, see "In Search Mode," by Dell Inc., June 2007, DELL.COM/downloads/global/casestudies/508_2007_google_v11.pdf; "Zoom-Zooming Ahead," by Dell Inc., September 2007, DELL.COM/downloads/global/casestudies/2007_Mazda_79990886_v3.pdf; and "Taking Stock," by Dell Inc., September 2007, DELL.COM/downloads/global/casestudies/2007_Unilever_79982804_v3.pdf.

VIRTUALIZING MICROSOFT EXCHANGE SERVER 2007 WITH VMWARE ESX SERVER 3

BY TODD MUIRHEAD
KONG YANG

Before virtualizing Microsoft® Exchange Server 2007, organizations should understand how those virtual machines might perform. This article describes tests used to evaluate the performance and scalability of Microsoft Exchange Server 2007 on a VMware® ESX Server 3 virtualization platform.

Related Categories:

Dell Enterprise Technology Center

Microsoft Exchange Server 2007

Virtualization

VMware

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

Initial enterprise virtualization deployments often focus on test and development and server consolidation. Once these initial deployments are successful, the next step is to identify additional applications that can be virtualized, often beginning with basic, noncritical applications. Finally, enterprises might consider virtualizing their most critical, performance-intensive applications—including Microsoft Exchange, a common application that can be critical to enterprise operations and requires high performance. Evaluating Exchange in virtualized environments can help determine whether other performance-intensive applications are suitable candidates for virtualization.[1]

A previous *Dell Power Solutions* article[2] examined how Microsoft Exchange Server 2003 performed on a VMware ESX Server virtualization platform running on a Dell™ PowerEdge™ 6850 server. Exchange Server 2007, however, introduces many different server roles and features, and as a 64-bit application, it offers higher memory scaling than Exchange Server 2003. To help demonstrate how Exchange Server 2007 performs in a virtualized environment, in August 2007 Dell engineers performed tests designed to address four key aspects of virtual machine (VM) sizing and performance: initial VM sizing, quality of service, appropriate resource levels for different user loads, and whether a single large VM or multiple small VMs provide the highest performance. The results indicated that in the test environment, appropriately configured VMs could support up to 500, 1,000, or 2,000 heavy Exchange users while continuing to provide acceptable quality of service.

Hardware and software configuration in the test environment

To provide an appropriate baseline VM configuration, the test team based the test environment on Dell reference architectures for Exchange Server 2007 and recommendations of the Dell Exchange 2007 Advisor tool.[3] This environment utilized a Dell PowerEdge 2950 server—a 2U, two-socket rack server that supports Intel® Xeon® 5100 and 5300 series processors and up to 32 GB of RAM, and provides a PCI Express (PCIe) riser with three PCIe slots or two PCI Extended (PCI-X) slots and one PCIe slot.

The test team configured this server with two quad-core Intel Xeon E5345 processors at 2.33 GHz, 16 GB of RAM, and VMware ESX Server 3.0.1. Each VM ran the Microsoft Windows Server® 2003 Enterprise x64 Edition OS and Microsoft Exchange Server 2007. All of the Exchange Server 2007 server roles were virtualized.

[1] When considering the deployment of Microsoft Exchange Server 2007 on a virtualization platform such as VMware ESX Server, please refer to support.microsoft.com/kb/897615 for information about the Microsoft support policy for their applications running in non-Microsoft virtualization software.

[2] "Microsoft Exchange Server 2003 Performance on VMware ESX Server 3," by Kong L. Yang and Aravind Pavuluri, in *Dell Power Solutions*, August 2007, DELL.COM/downloads/global/power/ps3q07-20070369-Yang-OE.pdf.

[3] See "Dell Exchange 2007 Advisor and Representative Deployments," by Farrukh Noman and Bharath Vasudevan, Dell Product Group, April 2007, DELL.COM/downloads/global/solutions/Dell_Exchange_2007_Advisor.pdf; and the Dell Exchange 2007 Advisor tool at DELL.COM/content/topics/global.aspx/tools/advisors/exchange_advisor.

The server was connected to a Dell/EMC CX3-80 storage array through McDATA Sphereon 4700 Fibre Channel switches and to client systems through a Dell PowerConnect™ 5324 switch. The client systems included a Microsoft Active Directory® domain controller, a Domain Name System (DNS) server, a VMware VirtualCenter server, and a Microsoft Exchange Load Generator (LoadGen) control server.

The Dell reference architectures for Exchange Server 2007 include small, medium, and large configurations. Figure 1 outlines the basic configurations for each VM. To help evaluate the performance of Exchange Server 2007 in a virtualized environment, the test team ran these small, medium, and large VM configurations in different tests; by examining key metrics during these tests, the team determined the configurations that delivered the highest performance while maintaining good quality of service.

The primary metrics used to compare performance were the host server's processor utilization and the 95th-percentile response time for the LoadGen SendMail task. The test team used Microsoft Windows® Performance Monitor (perfmon) to monitor VM performance and esxtop to monitor host server performance. The team configured perfmon to log processor, memory, disk, network, and Exchange-specific counters, which were logged in a comma-separated values (CSV) performance log file and later imported into a spreadsheet for analysis. The team configured esxtop to log processor, memory, disk, and network counters for both ESX Server and the VMs.

The test team used the LoadGen tool running on a client server to simulate a messaging workload utilizing the included heavy workload profile. LoadGen models the normal daily e-mail usage profile of real users and can help estimate the number of users that a system can support. It reports performance metrics that indicate how an Exchange server would typically respond when supporting a given number of users. LoadGen reports results in terms of response time or latency across 16 common e-mail tasks, including sending messages, deleting messages, browsing a calendar, logging on, and logging off. Although LoadGen tests do not necessarily match real-world deployments, they can provide usable data for comparison between different workloads. In the SendMail task chosen by the test team to evaluate Exchange performance, good quality of service is defined as the 95th percentile's average latency being less than 500 milliseconds.

	Small	Medium	Large
Users	500	1,000	2,000
Virtual processors	1–4	1–4	1–4
Memory	8 GB	8 GB	16 GB
Data disks	4	8	16
Log disks	2	4	8
Virtual network interface cards	1	1	1
Storage groups	4	8	16

Figure 1. *Basic virtual machine configurations in the test environment*

Test results: Response times and processor utilization

By taking advantage of the Dell reference architectures for Exchange Server 2007, the test team determined that the disk, memory, and network resources were sufficient at each configuration size. Therefore, the testing focused on analyzing the effects of virtual processors (vCPUs) on the overall performance of the Exchange VMs. The results across the majority of test cases showed good response times, with the 95th percentile for the SendMail task below or near 500 milliseconds. The response times only greatly exceeded this threshold in configurations where the processors or amount of memory were clearly inadequate.

Scaling up: Increasing virtual processors in one virtual machine

The test results indicated that the optimal number of vCPUs per VM is determined by the number of users that VM needs to support. When the VM was supporting 500 heavy users, one vCPU provided lower response times and more efficient processor utilization than two or four vCPUs (see Figure 2). When the VM was supporting 1,000 heavy users, two vCPUs provided the lowest response times (see Figure 3). When the VM was supporting 2,000 heavy users, four vCPUs provided the lowest response times (see Figure 4). In the 2,000-user tests, one vCPU was insufficient to complete the LoadGen tasks. Overall, the host server's processor utilization remained low—less than 25 percent—across all tests using all three types of VMs.

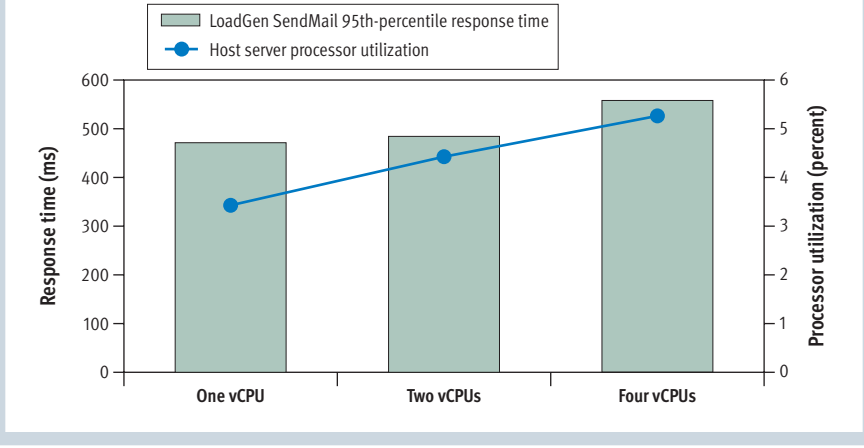


Figure 2. *Scaling up: Different numbers of virtual processors in one small virtual machine supporting 500 heavy LoadGen users*

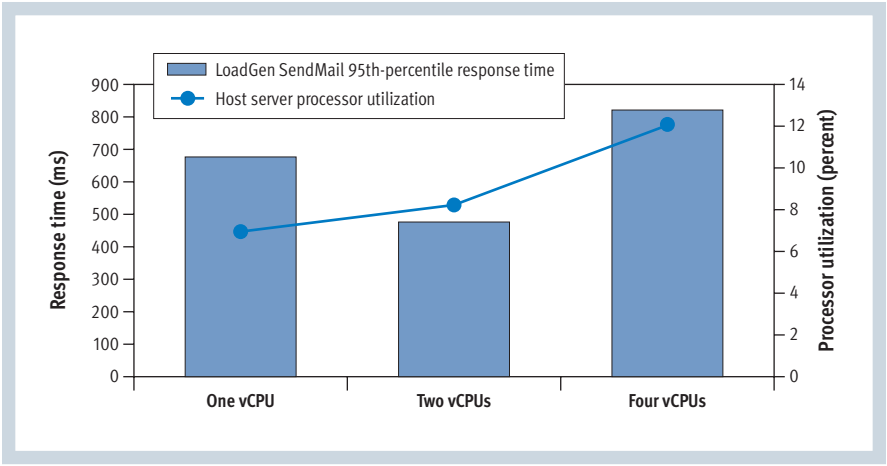


Figure 3. *Scaling up: Different numbers of virtual processors in one medium virtual machine supporting 1,000 heavy LoadGen users*

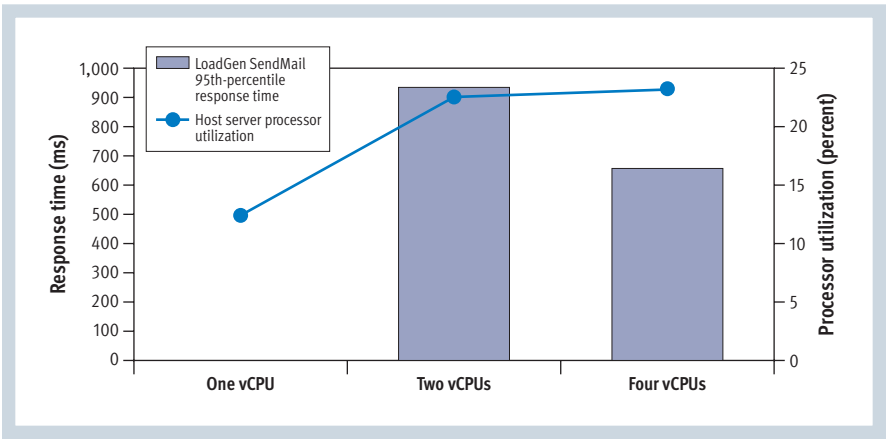
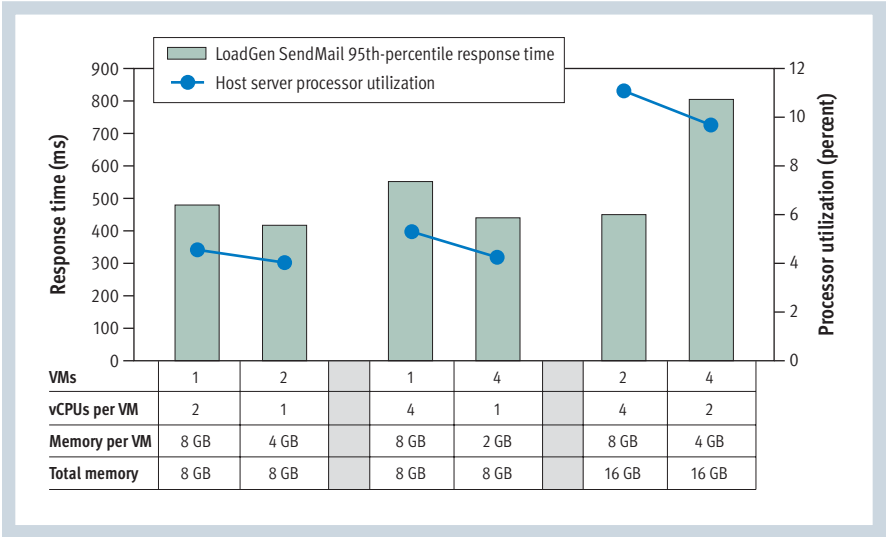


Figure 4. *Scaling up: Different numbers of virtual processors in one large virtual machine supporting 2,000 heavy LoadGen users*



VMs	1	2		1	4		2	4
vCPUs per VM	2	1		4	1		4	2
Memory per VM	8 GB	4 GB		8 GB	2 GB		8 GB	4 GB
Total memory	8 GB	8 GB		8 GB	8 GB		16 GB	16 GB

Figure 5. *Scaling out: Different numbers and configurations of small virtual machines supporting 500 heavy LoadGen users*

Scaling out: Varying virtual machines, virtual processors, and memory

In addition to examining how the VMs scaled up in different configurations with different numbers of vCPUs, the test team also evaluated how these systems scaled out with different combinations of VMs, vCPUs, and memory while keeping the disk and network subsystems constant. Figures 5–7 show the results. Providing sufficient processor and memory resources to the VMs helped reduce overall host server processor utilization and response times. However, assigning processor and memory resources beyond what was necessary tended to degrade performance.

Evaluating qualitative response times

As the preceding results show, although there were a few outliers in cases where the processor resources were insufficient for a particular user load, the tested configurations generally produced acceptable response times, including at least one set of LoadGen SendMail 95th-percentile response times that was near or below the 500-millisecond threshold. To help qualitatively evaluate the difference between a 95th percentile of 500 milliseconds and a 95th percentile of 1 second, the test team opened Microsoft Office Outlook® 2007 client sessions on two terminals, then composed and sent e-mail requests while LoadGen was

TALK BACK

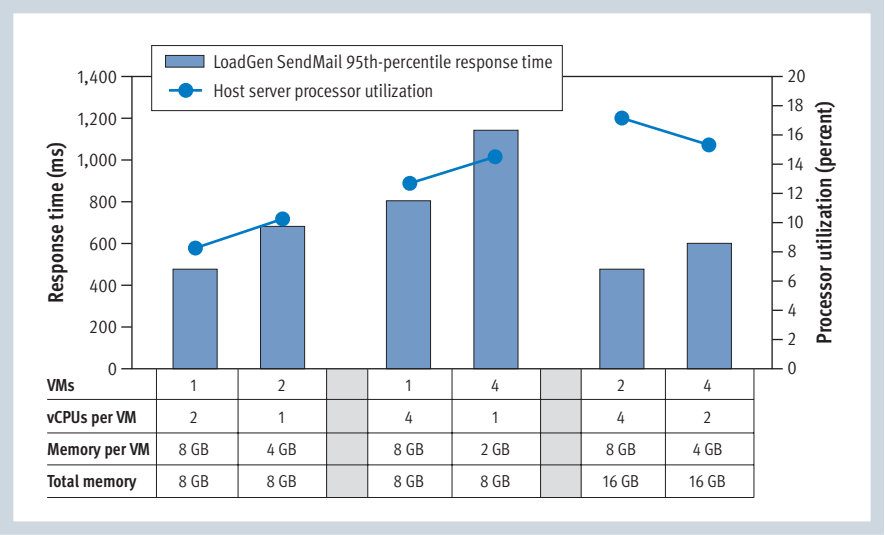
Tell us how the Dell Enterprise Technology Center can help your organization better simplify, utilize, and scale enterprise solutions and platforms. Send your feedback and ideas to enterprise_techcenter@DELL.COM.

running. In each case, as soon as a tester sent a message, the message went to the out-box and was in transit in less than 1 second. A few seconds later, the message appeared in the receiver's in-box. The difference between configurations at a 500-millisecond threshold and those at a 1-second threshold was imperceptible to these testers.

Microsoft Exchange performance in virtualized environments

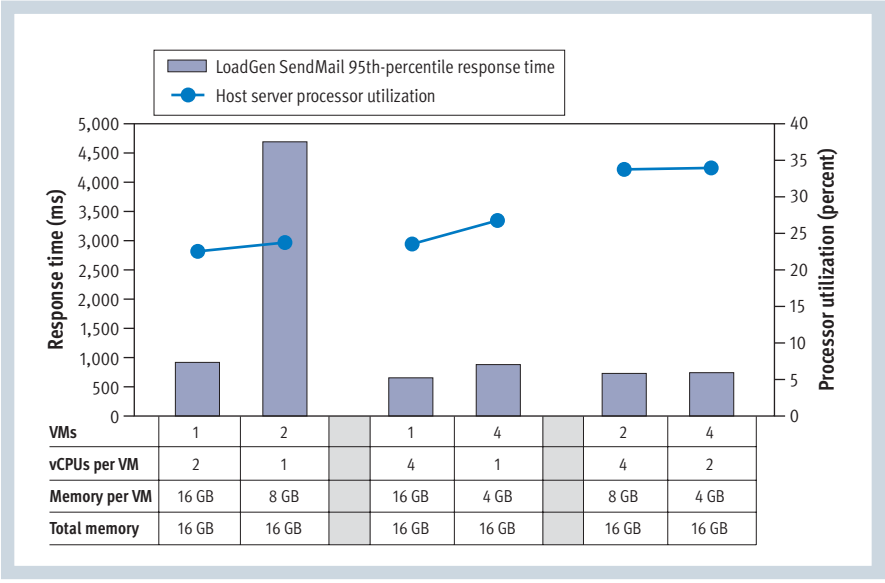
Understanding how Microsoft Exchange, a communication lifeline for many enterprises, performs in virtualized environments is key to planning data center deployments. In the Dell tests described in this article, a VM with one vCPU provided the lowest response times when supporting 500 heavy users, a VM with two vCPUs provided the lowest response times when supporting 1,000 heavy users, and a VM with four vCPUs provided the lowest response times when supporting 2,000 heavy users.

In these tests, the configurations with many small VMs provided performance and response times similar to those with a few large VMs—for this reason, enterprises should base their VM configurations on factors such as software licensing costs, ease of management, and enterprise standards. These factors can be particularly important when utilizing ESX Server features such as VMware VMotion™ and VMware High Availability (VMware HA) technology, because downtime and migration time can affect overall quality of service for end users. Although LoadGen results are not necessarily representative of production deployments, and performance will vary depending on the specific infrastructure, Exchange configuration, and workload profile, these results can provide a reference point for this particular configuration using this specific workload profile.



VMs	1	2		1	4		2	4
vCPUs per VM	2	1		4	1		4	2
Memory per VM	8 GB	4 GB		8 GB	2 GB		8 GB	4 GB
Total memory	8 GB	8 GB		8 GB	8 GB		16 GB	16 GB

Figure 6. *Scaling out: Different numbers and configurations of medium virtual machines supporting 1,000 heavy LoadGen users*



VMs	1	2		1	4		2	4
vCPUs per VM	2	1		4	1		4	2
Memory per VM	16 GB	8 GB		16 GB	4 GB		8 GB	4 GB
Total memory	16 GB	16 GB		16 GB	16 GB		16 GB	16 GB

Figure 7. *Scaling out: Different numbers and configurations of large virtual machines supporting 2,000 heavy LoadGen users*

Todd Muirhead *is a senior engineering consultant at the Dell Enterprise Technology Center.*

Kong Yang *is a systems engineer in the Dell Virtualization Solutions Engineering Group.*

MORE ONLINE
DELL.COM/PowerSolutions

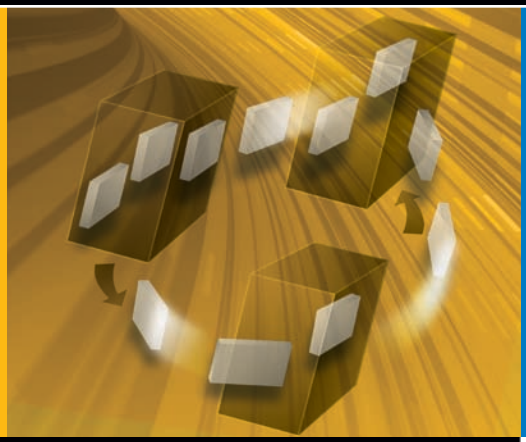
QUICK LINKS

Dell Enterprise Technology Center:
DELL.COM/TechCenter

Dell Enterprise Technology Center wiki:
www.delltechcenter.com

Dell and VMware:
DELL.COM/VMware

Microsoft Exchange solutions on Dell:
DELL.COM/Exchange



REDUCING POWER COSTS BY MIGRATING LEGACY SERVERS TO A VIRTUALIZED DELL POWEREDGE 2950

BY SCOTT HANSON

Virtualizing legacy servers nearing the end of their life cycle can help reduce power costs while maintaining similar levels of performance. Recent Dell tests show that migrating legacy servers to a virtualized Dell™ PowerEdge™ 2950 can help reduce average power use by approximately 2 kW, potentially saving more than US$6,000 over a three-year period.

Related Categories:

Dell Enterprise Technology Center

Dell PowerEdge servers

Power and cooling

Server consolidation

Virtualization

VMware

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

Server consolidation is a key advantage of virtualization in enterprise data centers. By migrating physical systems to virtual machines (VMs) on a single server, enterprises can efficiently utilize hardware resources and help reduce power costs while simplifying management for IT staff.

As legacy servers reach the end of their life cycle, enterprises should consider consolidating these systems as VMs on virtualized servers rather than replacing them with new physical servers. To help demonstrate the advantages of this approach, in July 2007 Dell engineers set up three legacy servers—a Dell PowerEdge 2650, IBM® eServer xSeries 345, and HP ProLiant DL380 G3—and measured their performance and power use. They then converted these systems to VMs on a Dell PowerEdge 2950 server running VMware® Infrastructure 3, increased the number of VMs to maximize resource utilization, and compared the power use of this virtualized server with the average power use of a comparable number of legacy physical servers. The results show that, on average, such a consolidation could help reduce power use by approximately 2 kW, and that consolidating multiple HP ProLiant DL380 G3 servers could provide savings of more than US$6,000 in power costs over a three-year period.

Hardware and software configuration in the test environment

The tests focused on consolidating approximately three-year-old Dell PowerEdge 2650, IBM eServer xSeries 345, and HP ProLiant DL380 G3 servers onto a Dell PowerEdge 2950 server running VMware Infrastructure 3 virtualization software. The hardware was chosen to be representative of servers that may be nearing the end of their life cycle, to help demonstrate how their existing workloads can be easily and efficiently migrated to a current-generation Dell PowerEdge server.

Figure 1 shows the hardware configuration for each server in the test environment. The test team disabled Intel® Hyper-Threading Technology on all of the servers, and downloaded and installed the latest firmware and driver versions available on the vendor Web sites. The hard drives in each server were configured in a RAID-5 array with no hot spare.

As shown in Figure 1, the test team set up each of the three legacy servers as similarly as possible: each was configured with two Intel Xeon® processors DP at 3.06 GHz with a 1 MB cache and a 533 MHz frontside bus, and had 2 GB of RAM. Each of these servers ran the Microsoft® Windows® 2000 Server OS with Service Pack 4 (SP4). The Dell PowerEdge 2950 used for consolidation was configured with two quad-core Intel Xeon X5355 processors at 2.66 GHz with an 8 MB level 2 (L2) cache and a 1,333 MHz frontside bus, and had 12 GB of RAM.

To help evaluate server performance and scalability, the test team used the Dell DVD Store Version 2 (DS2) e-commerce test application, which functions as a database test or as a stress tool. DS2 includes a back-end database component, a Web application layer, and driver programs, and includes advanced database features such as transactions, stored

procedures, triggers, and referential integrity. The test team used the Microsoft SQL Server™ 2000 Enterprise Edition database platform with SP4 to run the DS2 application. DS2 is freely available under the GNU General Public license at linux.dell.com/dvdstore. For more information, visit www.delltechcenter.com/page/DVD+Store.

To help ensure that the software stack remained consistent between the physical servers and VMs, the test team used VMware Converter 3 to migrate the servers to VM images. This robust, enterprise-class migration tool can be run on a wide variety of hardware and supports many common Microsoft Windows operating systems, enabling administrators to easily convert local and remote physical servers, third-party VM formats (such as Microsoft Virtual Server), and physical server backup image formats (such as Symantec Backup Exec System Recovery) to VMware VMs and perform other conversion and recovery tasks. A demonstration of the migration process used for the Dell tests is available at www.delltechcenter.com/page/Demonstrations.

Test results: Power use and cost savings

The Dell tests comprised three parts: testing the physical servers, testing the VMs, and calculating power use and cost savings.

Testing the physical servers

To test the physical servers, the test team used DS2 to increase the workload on each server until the processor utilization was approximately 30 percent. They then varied the DS2 parameters across multiple runs until they found a set of parameters that provided comparable processor utilization on each server: in this case, two threads and a think time of 0.015 seconds. Finally, they ran the DS2 application three times (for one hour each time) on each physical server and averaged the results, recording performance (in orders per minute [OPM]), power use (in watts), and average processor utilization. Figure 2 summarizes the results.

Although the purpose of these tests was not specifically to compare the physical servers, it is interesting to note that the Dell PowerEdge 2650 provided the highest performance of the three, while the HP ProLiant DL380 G3 had the highest power consumption.

Testing the virtual machines

The test team next migrated each physical server to a VM on the Dell PowerEdge 2950 using VMware Converter 3, then used the average OPM across the three physical servers as the baseline performance that the VMs had to achieve. After varying the DS2 parameters across several test runs, they determined that using two threads and a think time of 0 seconds yielded approximately 2,800 OPM from each VM. Because the migration to a virtualized environment eliminated hardware-based performance differences between the three systems,

	Dell PowerEdge 2650	IBM eServer xSeries 345	HP ProLiant DL380 G3	Dell PowerEdge 2950
Processors	Two Intel Xeon processors DP at 3.06 GHz	Two Intel Xeon processors DP at 3.06 GHz	Two Intel Xeon processors DP at 3.06 GHz	Two quad-core Intel Xeon X5355 processors at 2.66 GHz
Processor cache	1 MB	1 MB	1 MB	8 MB
Frontside bus speed	533 MHz	533 MHz	533 MHz	1,333 MHz
Memory	2 GB (four 512 MB PC2100 dual in-line memory modules [DIMMs])	2 GB (four 512 MB PC2100 DIMMs)	2 GB (four 512 MB PC2100 DIMMs)	12 GB (six 2 GB PC5300 DIMMs)
RAID controller	PowerEdge Expandable RAID Controller (PERC) 3/Di	IBM ServeRAID-5i	HP Smart Array 5i+	PERC 5/i
Internal disks	Three 73 GB, 10,000 rpm, 3.5-inch SCSI drives	Three 73 GB, 10,000 rpm, 3.5-inch SCSI drives	Three 73 GB, 10,000 rpm, 3.5-inch SCSI drives	Five 73 GB, 10,000 rpm, 2.5-inch Serial Attached SCSI (SAS) drives
Power supplies	Dual redundant	Dual redundant	Dual redundant	Dual redundant

Figure 1. *Hardware configuration in the test environment*

	Performance	Power use	Processor utilization
Dell PowerEdge 2650	3,041.33 OPM	290.67 W	30.82%
IBM eServer xSeries 345	2,450.67 OPM	278.67 W	28.91%
HP ProLiant DL380 G3	2,858.00 OPM	346.00 W	33.20%
Average	**2,783.33 OPM**	**305.11 W**	**30.98%**

Figure 2. *Average performance, power use, and processor utilization results for the physical servers*

TALK BACK

Tell us how the Dell Enterprise Technology Center can help your organization better simplify, utilize, and scale enterprise solutions and platforms. Send your feedback and ideas to enterprise_techcenter@DELL.COM.

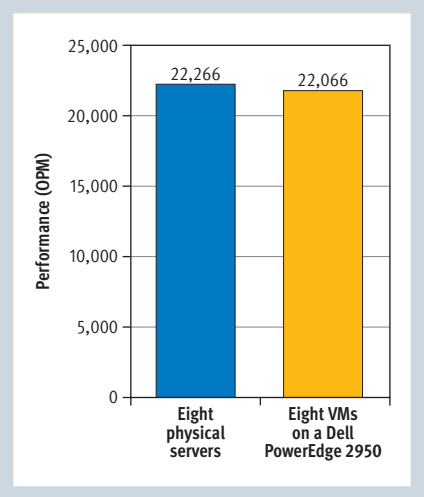


Figure 3. *Average performance extrapolated for eight physical servers compared with the performance of eight virtual machines on a Dell PowerEdge 2950*

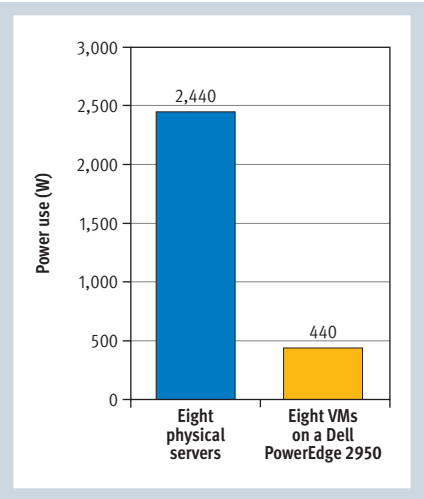


Figure 4. *Average power use extrapolated for eight physical servers compared with the power use of a single virtualized Dell PowerEdge 2950*

the three VMs were essentially identical, and running the DS2 application on each of them yielded the same OPM results.

The test team then increased the number of VMs running on the PowerEdge 2950 until the processor utilization measured in VMware ESX Server was approximately 80 percent, which resulted in the consolidation of eight workloads on this server. Extrapolating the average performance of the physical servers shows that eight of these servers could process approximately 22,266 OPM. When the test team ran the DS2 application on the eight VMs, they processed a total of 22,066 OPM, which is within 1 percent of the projected average performance of the physical servers (see Figure 3).

In terms of power, extrapolating the average power use of the physical servers shows that eight physical servers would use approximately 2,440 W. The virtualized Dell PowerEdge 2950, in contrast, used only 440 W—a savings of 2 kW, resulting in a system more than five times more energy efficient than eight of these physical servers (see Figure 4). Comparing the virtualized Dell PowerEdge 2950 with eight HP ProLiant DL380 G3 servers, which would have the highest power use of the three servers, at 2,858 W, shows that the virtualized PowerEdge 2950 would be more than six times more energy efficient than eight of these physical servers.

Calculating power use and cost savings

Reducing power use is not just an abstract benefit—controlling power costs can be critical to building an efficient IT infrastructure. In the test environment, the virtualized Dell PowerEdge 2950 saved 2,418 W compared with what eight HP ProLiant DL380 G3 servers would consume. Power is sold by the kilowatt-hour (kWh). Assuming that these servers would remain in use for three years and run constantly during that time, the total difference in power use between the virtualized PowerEdge 2950 and eight HP ProLiant DL380 G3 servers over that period would be 63,545.04 kWh.

The average price for commercial electricity in the United States as of June 2007 was US$0.0992/kWh.[1] Multiplying this price by the three-year difference in power use gives a total cost savings over that period of US$6,303.67.

Efficient consolidation to a virtualized Dell PowerEdge server

Using a tool such as VMware Converter 3, administrators can easily migrate physical servers to VMs without disrupting the existing environment. As demonstrated by the Dell tests, consolidating eight of the legacy servers described in this article to a Dell PowerEdge 2950 server running VMware Infrastructure 3 could help reduce power consumption by approximately 2 kW; based on these results, consolidating eight HP ProLiant DL380 G3 servers could save US$6,303.67 in power costs over a three-year period. Importantly, enterprises can take advantage of these savings while still maintaining similar levels of performance. As virtualization becomes increasingly mainstream, this type of consolidation can provide an efficient way to reduce power costs for enterprises with physical servers nearing the end of their life cycle.

Scott Hanson *is a senior engineering consultant on the Dell Enterprise Technology Center team. Scott has a B.S. in Computer Science from the University of North Texas and holds Microsoft Certified Systems Engineer + Internet (MCSE+I) and Red Hat® Certified Engineer (RHCE) certifications.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell Enterprise Technology Center:
DELL.COM/TechCenter

Dell Enterprise Technology Center wiki:
www.delltechcenter.com

Demonstration of VMware Converter 3 conversion process:
www.delltechcenter.com/page/Demonstrations

[1] "Average Retail Price of Electricity to Ultimate Customers by End-Use Sector, by State, June 2007 and 2006," by the Energy Information Administration, U.S. Department of Energy, in *Electric Power Monthly*, September 2007, www.eia.doe.gov/cneaf/electricity/epm/table5_6_a.html.

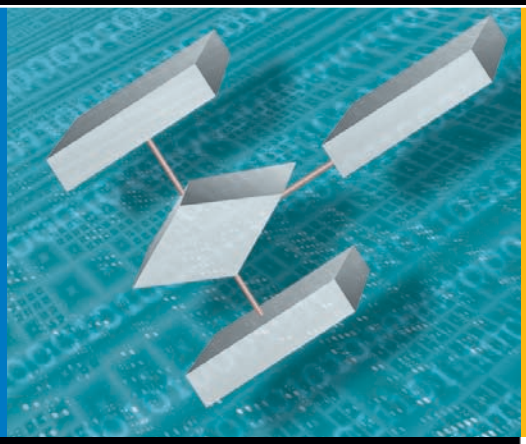
Sometimes, having the best cards has nothing to do with luck.

ATI FIREGL GRAPHICS

Don't gamble with your graphics. Optimized and certified for CAD and DCC applications, the all new ATI FireGL™ 3D graphics accelerators from AMD are your essential workstation graphics cards—delivering increased performance, reliability, and value.

AMD
Smarter Choice

Featured in Dell Precision™ Workstations. Visit ati.amd.com/FireGL for more information.



DELL IT IMPLEMENTS FACTORY AUTOMATION AND VISIBILITY IN REAL TIME

BY PETER SWARTZ
CLAY JOHNSON
RICHARD BROWN

At Dell, maintaining high efficiency in supply chain and manufacturing operations is key to success. So when Dell IT saw the number of custom applications used in its factories spiraling out of control, it responded by planning and implementing a single standardized application layer, while simultaneously integrating tools allowing mobile users to easily access and control factory operations in real time.

Related Categories:

Application development

Best practices

Manufacturing automation

Visit ***DELL.COM/PowerSolutions*** *for the complete category index.*

Because Dell is known for running an efficient supply chain, it may seem that the company has already fully tapped this particular source of process improvement. But, on the contrary, Dell continues to identify areas in which innovative use of technology can further streamline its supply chain and manufacturing operations while continuing to increase business manageability and control.

For example, although Dell is a global company, it is always looking for ways to become more *globalized* in its operations. Over the past several years, Dell expanded its manufacturing and distribution capabilities across the United States and into Asia, Europe, and South America to help keep pace with worldwide demand. Although each of these facilities shares a core set of application instances, each factory rollout required the local development of several custom applications to solve specific problems. Over time, the number of these applications grew to several hundred globally, creating unnecessary complexity—and an opportunity for Dell IT to help the company gain additional competitive advantage.

Factory automation initiative helps Dell go global

As part of a sweeping effort currently underway to shape Dell IT into a more global organization (with a global application base and single global owners for every process and shared service, rather than regional IT groups supporting regional applications), Dell performed extensive analyses of manufacturing software and selected best-of-breed applications for global deployment.

One of the first areas addressed was factory control. The selected applications performed their specific jobs well, but none of them were designed to directly communicate with or control the production line hardware. Consequently, each location had to develop a different application to connect IT systems with this hardware to provide physical control over the production line. Most of these applications had been developed by an automation provider and only focused on a small area of the factory, and the integration was typically crude.

But although the production line hardware differed among various facilities, the underlying business rules did not. For example, say a factory needs to direct a product of type A to a work cell that can process or build that type of product. The business rules that determine whether a product is type A may be complex, but they do not change. However, the work cells that can process or build type A products are dynamic; they can change in real time, or may differ between facilities. To help fully globalize the workflow, the details behind directing a type A product to a type A work cell must be abstracted, with the *how* and the *why* separated to avoid the need to customize every instance.

Dell already had a small set of factory applications of this type, created with off-the-shelf Proficy HMI/SCADA

CIMPLICITY software from GE Fanuc. CIMPLICITY is designed to help develop supervisory control and data acquisition (SCADA) applications—a broad category that includes factory automation. The Dell factory applications were both highly stable and available, and augmented by in-house development knowledge and ready access to source code.

CIMPLICITY is in widespread use in several industries, from energy and water suppliers to large automobile manufacturers. Many of the factory lines that use SCADA products, such as those of automobile manufacturers, produce a low mix of complex products on a relatively low-speed production line. Other companies produce a single, simple product, such as an injection-molded plastic toy, on a high-speed production line. For these manufacturers, the majority of the decision logic and control over the line resides within the SCADA application rather than in other systems, and they use SCADA to provide "closed-loop" control.

Dell, in contrast, builds a high mix of complex, customized products on a high-speed production line. These products are relatively simple to assemble, but building customized PCs with the exact hardware, software, and accessories that a customer orders—and not until he or she orders it—is an intricate process that is quite different from other manufacturing environments. While other manufacturers can configure most of the business logic for routing and tracking material into their factory control applications with little interaction or dependency on external systems, Dell relies on specialized applications to determine, in real time, what needs to happen with a particular order.

Consequently, Dell needs a measure of closed-loop control, but the business logic is too complex to reside completely within the SCADA layer. CIMPLICITY had the components to meet the challenge, given the right level of deployment innovation and expertise.

Because of these advantages, in late 2005 Dell IT elected to standardize on CIMPLICITY and Proficy Tracker software to build its global factory control integration layer. Dell was already designing a new factory, the timing of which meshed well with Dell IT plans to create this layer to integrate its global manufacturing applications with the factory line hardware and enable an efficient, manageable production line. The production line would be fully automated, from the initial kitting operation through to distribution, and would provide comprehensive routing and tracking functionality for every product moving down the line.

To accomplish this integration, Dell IT built an additional layer using the Microsoft® .NET Framework to provide a common, simplified interface between the routing and tracking application and other factory systems (see Figure 1). The project required a high degree of collaboration between Dell engineering and application teams to help ensure that the production line would operate as expected and sustain the necessary rates. Also key to success was the inclusion of strong program and project management functions as well as alignment with business owners to help ensure that expectations and timelines were met. In this particular case, the Dell Factory Operations organization provided key requirements for how the factory needed to operate, while Dell Factory Controls turned those requirements into standards and then implemented those standards on the hardware controls side. Dell also worked closely with its major technology partners, including Microsoft and GE Fanuc, to help maximize the system's capabilities.

Integration layer boosts management efficiency

Advanced product routing and tracking through the manufacturing process has been highly advantageous for Dell in terms of both actual throughput and effective factory monitoring and management. Dell can gather many production metrics—including units per hour, total units produced, cycle time, and work-in-progress levels in different areas—in real time, and with much greater detail and accuracy than was previously possible.

The Tracker layer in the application stack interfaces with several other factory applications, enabling it to monitor key metrics related these interfaces. For example, it can gather data on response times and the number of call failures, and report on the overall health of factory systems. Similarly, Tracker applications interface with the production line controllers to report on equipment performance, efficiency, and current hardware issues. With this data

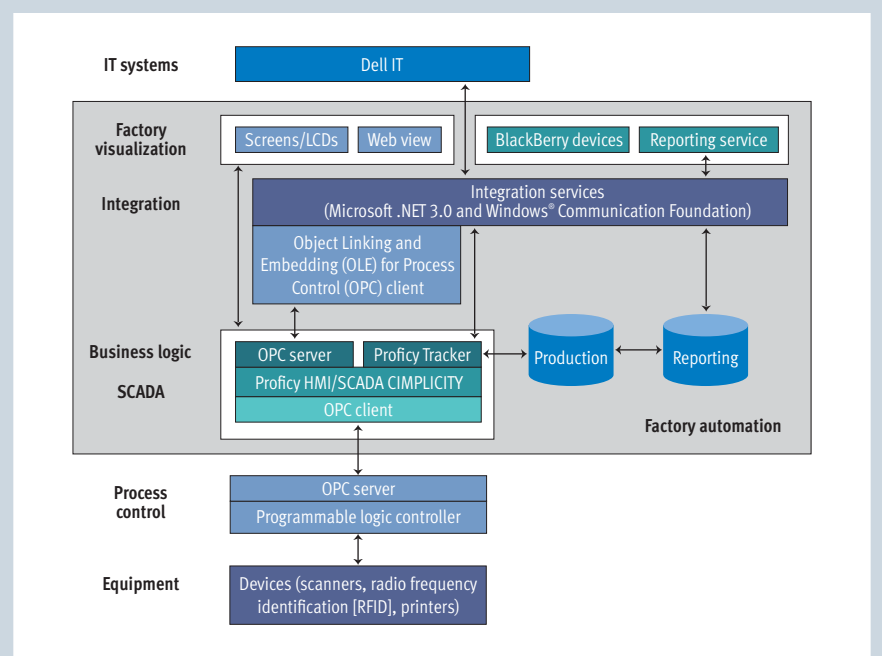


Figure 1. *Dell factory automation architecture*

now available in real time, operations personnel can monitor the current status and react to problems as they occur.

Because Tracker applications function as an integration layer, they must interface with the production line hardware on one side and the various Dell IT manufacturing systems on the other. On the hardware controls side, they interface with a variety of controllers—which presented difficulties during implementation, because Dell had contracted different factory systems integrators to write control logic at different times, creating a lack of standardization at the controls layer that Dell IT had to anticipate and manage. On the factory application side, even though the Dell globalization efforts have helped dramatically reduce the total number of factory applications, a large number of systems remain. Each of these systems control or monitor some aspect of the Dell manufacturing or shipping process, and the Tracker applications must interface with them.

Dealing with these challenges required standardization—including a standard method of communicating between the various layers. On the hardware controls side, standardization began with the physical hardware devices and the interface to these devices. Dell IT defined a handshake protocol to help provide a standard and reliable method for passing information between the hardware controls and the integration layer. Because most factory systems had already begun to standardize on Web services, that technology became the major method of communication between the integration layer and other Dell IT systems.

Unifying the basic factory control systems on a single global platform was a step forward for Dell, but the project has not stopped there: Dell continues to refine and enhance a standard global code base for Tracker applications. This code base is used in new Dell factory implementations, and has also been rolled out to several existing facilities.

In a relatively short period of time, Dell IT took an off-the-shelf product and found innovative ways to use it, providing significant value to manufacturing operations. Compared with the previous environment, Dell manufacturing operations are more automated and efficient with the Tracker layer in place, while manufacturing operations staff have more data—data that is globally consistent—with which to manage the factory environment. Dell IT also took this opportunity to extend factory control and automation functionality down to the level of product routing and tracking through the factory based on complex business rules—in real time.

Dell IT extends factory visibility and mobile management

Before this project was underway, applications at the SCADA layer were typically funded and managed through the Dell Factory Operations and Factory Controls groups. However, these applications often focused only on a single section of a single factory. The globalization project, therefore, represents a new paradigm for how Dell approaches factory automation.

Creating this key application layer and making it a global standard was no small task. Dell augmented its own IT team with experienced developers and brought in expert consultants to help guide the team, provide additional staff, and drive knowledge transfer.

Perhaps the greatest challenge of the project was the fact that it radically changed the entire Dell factory automation control layer—the "circulatory system" of the factory. Regardless of how many orders are waiting, how many parts are available, how many team members are ready to build, and how many trucks are standing by to ship the orders out, nothing moves through the factory without the automation and control systems. When an enterprise is manufacturing more than 100,000 systems each day, even an hour of downtime can have a huge impact on revenue, efficiency, and customer satisfaction. Mitigating these risks required Dell to first deploy Tracker applications in a new factory where it could comprehensively test, troubleshoot, and refine system performance, helping ensure high levels of stability and availability once products began rolling down the line.

One of the most interesting results generated by the Tracker project was unanticipated. During planning and implementation, Dell IT focused on the potential benefits of creating a single global platform for factory automation and on tracking and controlling the factory lines—but the resulting visibility into real-time factory performance has turned out to have significant advantages. Because a common application layer now connects to the factory floor from end to end, administrators can monitor total factory output, current throughput, and work-in-progress levels—in real time, and at a level of detail unavailable to Dell in the past. This detail gives the factory operations staff increased visibility into their operations, allowing them to quickly detect and adjust to operational problems.

This real-time visibility has already spawned additional projects at Dell. One of these projects, for example, provides real-time monitoring of factory equipment, allowing technicians to be dispatched to correct hardware issues quickly—often before workers on the factory floor are aware of a looming problem.

Another innovative aspect of the factory automation project is its ability to globalize the company through information mobility and application access. Being a global company means that Dell employees often travel between sites, or may work from home or other remote locations. To help provide users with the same level of information and access regardless of

"Because of the critical nature of its work, the Dell IT group is a key part of innovation at Dell, in both the technologies the company depends on and the underlying business processes."

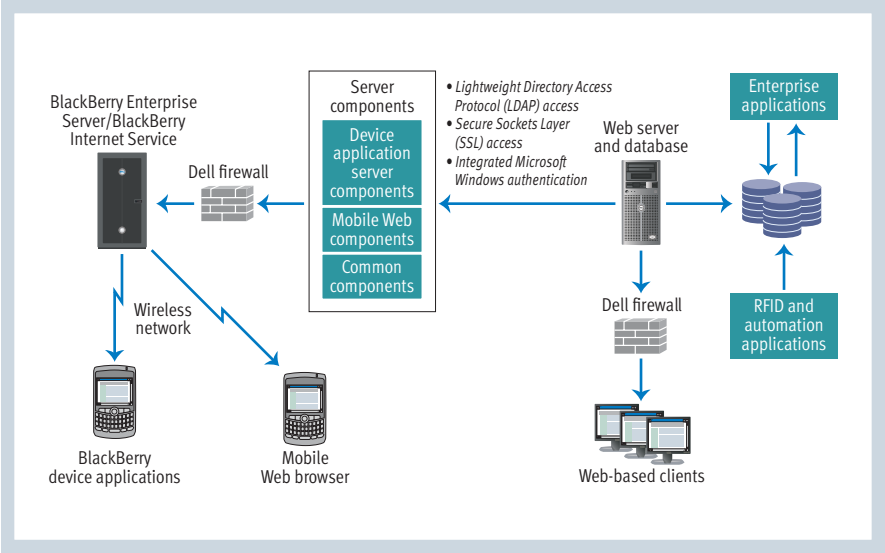


Figure 2. *Dell application architecture for BlackBerry device users*

location, Dell IT incorporated mobile platforms into the Tracker applications (see Figure 2). Working with Research In Motion (RIM) and Cingular Wireless (now AT&T), Dell significantly increased access flexibility by providing BlackBerry® device users with the same production data available internally.

In the next phase of this project, Dell plans to extend the mobile capabilities from mobile information delivery to mobile application control by enabling BlackBerry device users to make operational and workflow decisions on their handheld devices that affect factory operations in real time. For example, a manufacturing operations executive traveling in Bangalore would be able to receive operational information from a Dell factory in Brazil or Ireland or North Carolina—and, with a few clicks of the scroll wheel, make real-time workflow decisions for those production lines.

Automation enhancements increase control and return on investment

Although global implementation of the Tracker application layer is not yet complete, demand for it is extremely high. Dell has calculated that the increased operational visibility alone could save US$3.7 million per year at a single Dell factory. Extrapolating these savings over an eventual total of 12 factories globally leads Dell to anticipate a substantial return on investment (ROI), which should continue to accrue as the remaining factories implement the global application and code base.

Other results of the project are difficult to quantify. For example, Dell has seen increases in overall system reliability: implementing a common layer in place of a large number of individual applications has helped dramatically reduce the number of application-related incidents. And because a single application support group can provide incident resolution, the mean time to repair has also decreased significantly.

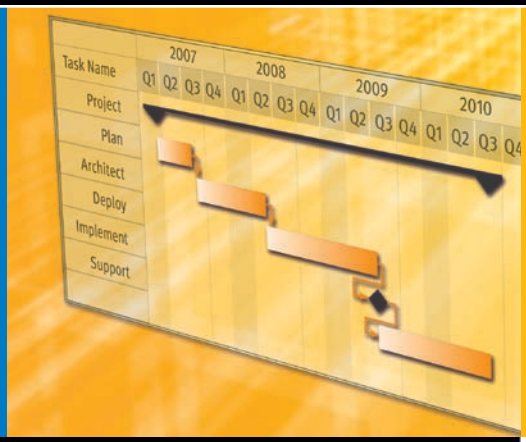
Increased control over the routing of material through the factory has also enabled several business process enhancements. Dell anticipates an annual ROI in excess of US$40 million related to these ancillary projects. And although these enhancements did not all directly result from the Tracker integration layer, they could not have been initiated without it.

Dell is not only in the business of IT; it uses IT every day to enable that business. Because of the critical nature of its work, the Dell IT group is a key part of innovation at Dell, in both the technologies the company depends on and the underlying business processes. By seizing an opportunity to enhance Dell manufacturing operations and then planning and implementing the technology to take advantage of that opportunity, Dell IT has helped significantly increase the efficiency and manageability of the company's core business.

Peter Swartz *is the director of IT for the Dell Global Manufacturing, Finance, and Supply Chain Group. He manages a US$250 million application portfolio on which Dell runs its worldwide supply chain, manufacturing and distribution plants, financial and corporate systems, and product development applications; has direct responsibility for the application team that develops and supports the real-time data and reporting for the Dell supply chain and factories; and is responsible for the IT aspects of new Dell manufacturing facilities worldwide. Peter has a degree in Electrical Engineering from the University of Pennsylvania and a degree in Economics from the Wharton School of the University of Pennsylvania.*

Clay Johnson*, IT manager for the Dell Global Manufacturing, Finance, and Supply Chain Group, is directly responsible for managing the automation software development team. He has day-to-day responsibility for developing, implementing, and sustaining automation solutions at Dell manufacturing, fulfillment, and distribution locations worldwide. Clay has a bachelor's degree in Computer Science Engineering and an M.B.A. from the University of Texas, and currently serves on the Engineering Industry Advisory Board for Southern Methodist University, the Technical Advisory Board for the University of North Texas, and the Manufacturing Peer Board for AMR Research.*

Richard Brown *is a development manager for the Dell Global Manufacturing, Finance, and Supply Chain Group. He manages a team of developers responsible for implementing and deploying automation and data acquisition solutions for Dell manufacturing and fulfillment centers worldwide. Richard has a degree in Computer Information Science from St. Edward's University and has spent 12 years in the field of high-tech manufacturing automation.*



DRIVING BUSINESS PROCESS INNOVATION WITH SERVICE-ORIENTED ARCHITECTURE AND SAP NETWEAVER

BY MICHAEL BECHAUF

Enterprise service-oriented architecture (SOA) can help reduce IT costs and increase flexibility, but successful deployment depends on the broad adoption of industry standards. SAP is committed to standards development and the integration of standards into the SAP® NetWeaver® platform, the foundation of SAP enterprise SOA.

Related Categories:

Enterprise Resource Planning (ERP)

SAP

Service-oriented architecture

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

Standards are essential to the efficient development and adoption of information management technologies—hardware, software, and services.[1] Accordingly, many commercially successful technologies have undergone some process of standardization before achieving mainstream adoption.

Consider the standards that make up the foundation of the World Wide Web, such as HTTP, HTML, and TCP/IP. These standards, and the cost-effective systems and software they allow, enabled the rapid development and success of the Internet, and were remarkably effective in making its technology available to millions of users. Conversely, a lack of standards can significantly impede broad adoption of IT. Before the World Wide Web, the lack of a standard network stack resulted in a highly fragmented market, where interoperability problems meant that many software applications could only be deployed on a specific LAN and that IT organizations were burdened with supporting disparate network infrastructures.

To help address these challenges and increase enterprise value, standards are a prerequisite for the widespread adoption of many technologies. Enterprise service-oriented architecture (SOA) is no exception. Openness—the foundation of enterprise SOA—and a holistic approach to standards are essential to SAP co-innovation with its partners and customers.

Utilizing standards as the foundation of enterprise SOA

Enterprise SOA is the blueprint for the latest generation of SAP software—the business applications and underlying technologies that form a *business process platform*. Enterprise SOA elevates the design, composition, and deployment of Web services to an enterprise level through the use of enterprise services, standards-based ways of encapsulating enterprise functionality and exposing it as a reusable business service. A combination of granular Web services and business logic, enterprise services form the building blocks needed to automate complex business processes and allow cost-efficient development of composite applications.

Enterprise SOA reverses the concept of integration. Traditionally, organizations purchased large enterprise systems and then spent tremendous amounts of capital on integration. With enterprise SOA, integration starts when it should—at the beginning. Organizations first acquire services—from SAP or SAP partners, for example, or by developing them internally—and then integrate them into composite applications to help address a specific set of challenges.

However, the business process flexibility enabled by these composite applications requires services from diverse sources, which is why standards—technical specifications

[1] A version of this article originally appeared in the "SAP Ecosystem in Action" special feature in *SAP Insider*, April–June 2007, www.sapinsideronline.com.

adopted to allow products from different sources to work together—are critical. For organizations using SAP software, standards allow both SAP and non-SAP applications and services to interoperate, meaning they can exchange information as seamlessly as possible without introducing unnecessary costs and risks.

Creating standards for technology, languages, and business semantics

SAP takes a holistic approach to the relationship between standards and the business process platform, which SAP classifies into three layers, as shown in Figure 1: technology standards, languages for defining business semantics, and business semantics standards. Utilizing standards within each layer allows the efficient assembly of interoperable composite applications. (For more information on the common standards that cut across all three of these layers, see the "Building common standards" sidebar in this article.)

Technology standards: Building security, reliability, and scalability

Technology standards provide the foundation for the openness and interoperability of the SAP enterprise SOA, which is designed to provide high levels of process and development flexibility. The technical underpinning of the SAP enterprise SOA is SAP NetWeaver, a comprehensive platform for the provisioning, production, consumption, and management of services and service-enabled business processes. This platform helps ensure that critical processes are secure, reliable, and scalable.

SAP NetWeaver is built on an enterprise SOA that utilizes Web services standards. These standards are important for efficiently deploying enterprise solutions, and are often created within international standards development organizations with business experts, systems architects, and other thought leaders. SAP leads and participates in these organizations and the development of these standards. For example, SAP is part of the World Wide Web Consortium (W3C) advisory board, which defines the technical standards for the World Wide Web, including HTML, XML, and core Web services specifications such as Simple Object Access Protocol (SOAP), Web Services Description Language (WSDL), Web Services Addressing (WS-Addressing), and Web Services Policy (WS-Policy).

SAP also leads multiple SOA standards critical to enterprise SOA at the Organization for the Advancement of Structured Information Standards (OASIS). These standards are reflected in the core Internet connectivity and Web services interoperability capabilities of SAP NetWeaver. This interoperability is also rigorously tested in the Web Services Interoperability Organization (WS-I), chaired by SAP. WS-I promotes consistent and reliable interoperability among Web services across platforms, operating systems, and programming languages. SAP NetWeaver is already compliant with WS-I Basic Profile 1.1 and WS-I Basic Security Profile 1.0.

Enterprise SOA is much more than just a technology platform, however: it integrates SOA with business semantics. Accordingly, the languages in which business semantics for services are defined must also be standardized.

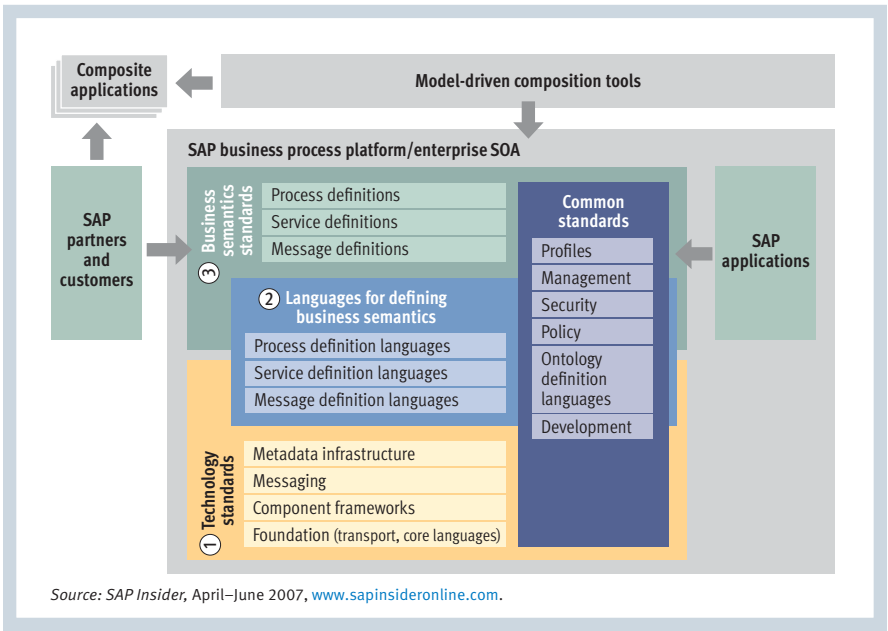
Languages for defining business semantics: Promoting a common vocabulary

For an organization to create composite applications using enterprise services, it needs a technology foundation of relevant Web services standards that help ensure secure, reliable message exchange. But it also requires common languages that can be understood by all parties when designing, provisioning, composing, and consuming enterprise services.

Such languages are used to create formal, standardized definitions of business information, processes, and services. For example, WSDL is a services definition language that provides a technology-independent way to describe Web services. These types of languages provide a bridge that allows technology and business semantics to evolve independently of one another. They also allow the expression of the business contract that defines the obligations of an enterprise services provider. Rigorous contract specification helps ensure the business integrity needed to develop and integrate composite applications.

Business semantics standards: Crossing industry borders

Business semantics standards enable the precise, efficient use of information across different technologies, markets, industries, and locations.



Source: SAP Insider, April–June 2007, www.sapinsideronline.com.

Figure 1. *Enterprise SOA standards*

BUILDING COMMON STANDARDS

Common standards apply to all layers of the standards taxonomy and are critical to a successful enterprise SOA strategy. SAP is leading the way in its adoption of industry-accepted common standards:

- The SAP NetWeaver development environment is built on the open source Eclipse platform. By using Eclipse, SAP has adopted the de facto industry-standard development environment, and organizations using SAP software can benefit from the large Eclipse ecosystem.
- SAP became a leader in enterprise Java when it announced early compatibility with Java Platform, Enterprise Edition 5 (Java EE 5).
- SAP is a key participant in the Open SOA Collaboration, providing developers with a simple, powerful way to construct composite applications. This collaboration defines a language-neutral model that exploits SOA characteristics and benefits. Currently, the collaboration is working on the Service Component Architecture (SCA) and Service Data Objects (SDO) specifications.
- SAP is also a key contributor to the Web Services Business Process Execution Language Extension for People (BPEL4People) specification, which provides missing process definition capabilities for human interactions as one of the key SOA building blocks.

For example, a purchase order can be correctly processed only if each field in the message is correctly understood in the proper context. Business semantics standards provide the common understanding necessary to execute a business process such as order-to-cash, which may include messages such as orders, ship notices, goods receipts, invoices, and remittances.

Today, many organizations are helping define which standards are necessary for specific vertical industries or for cross-sector uses. These organizations typically include private-sector entities seeking to establish a cooperative relationship with national, regional, and international standards organizations.

SAP, with its 30-year history of building business process applications, has helped lead the development of business semantics standards for vertical industries such as the aerospace and defense, automotive, chemical, consumer goods, high-tech, mill products, oil and gas, banking, health care, insurance, transportation, and public services industries. Today, SAP is actively engaged in over 50 vertical-industry standards development organizations as well as numerous customer focus groups and industry value networks.

The proliferation of vertical-industry standards organizations, however, has become a serious obstacle to interoperability, particularly when those standards cross industry borders. For example, functionally equivalent messages are represented in vastly different ways in different industries and sectors. Semantically, the messages have the same meaning and function, but integration efforts have become extraordinarily expensive across industries.

A convergence of methodologies and semantics across vertical industries could help address these challenges. Creating this convergence would require a well-positioned, cross-industry standards development organization. Accordingly, SAP has invested heavily in the United Nations Center for Trade Facilitation and Electronic Business (UN/CEFACT). UN/CEFACT standards define both basic message primitives (vocabulary) and the methodology that enables consistent naming and structuring (grammar). SAP adopted the UN/CEFACT methodology in enterprise services design several years ago, and has since been active in standards organizations in many vertical industries to encourage the adoption of this methodology.

Establishing common business semantics is an enormous investment, but success can facilitate the high level of interoperability crucial to enabling organizations to efficiently compose applications in an enterprise SOA. Industry standards and SAP enterprise services are already helping increase semantic interoperability as a result of these efforts.

Enhancing flexibility through industry standards

Industry standards defined by organizations such as W3C, OASIS, WS-I, and UN/CEFACT are critical to realizing the benefits of enterprise SOA. With SAP NetWeaver and enterprise SOA, SAP has enhanced business process flexibility so that organizations can develop, build, and manage business applications on a foundation of open standards—flexibility that, in turn, helps them use IT to increase the responsiveness and profitability of their business.

Michael Bechauf *is the vice president of industry standards at SAP, where he is responsible for SAP participation in industry-standards activities and for the strategy to integrate standards into SAP business process platform software. He currently serves as president and chairman of WS-I and as a member of the Java Community Process Executive Committee and the Eclipse Foundation board of directors.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

SAP:
www.sap.com

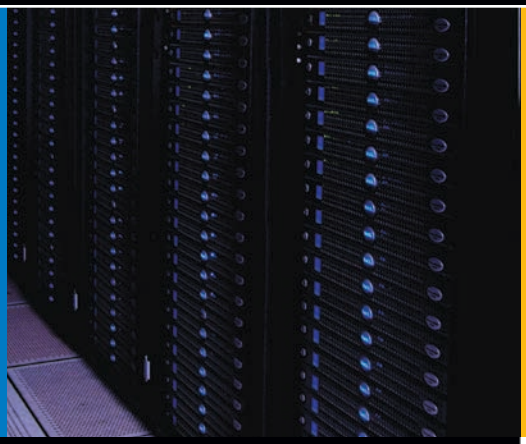
Standards and SAP enterprise SOA:
www.sdn.sap.com/irj/sdn/standards

EXPERT ADVICE.
PEER SUPPORT.
LAME JOKES.

How many experts does it take to solve a custom development problem? At sdn.sap.com, you'll find 850,000 developers, systems managers and other insiders to help with your toughest applications challenges and coding snafus. Not to mention free sample downloads, advice from SAP staff and maybe even a few new punch lines.

// JOIN IN AT SDN.SAP.COM

THE BEST-RUN BUSINESSES RUN SAP™



PACKAGE DEAL: HELPING TO REDUCE IT COMPLEXITY WITH DEDICATED APPLIANCES

BY JALEEL MOHAMMED
DAVID PENNINGTON

Dedicated appliances help enterprises deploy critical data center capabilities; streamline support, development, and testing; accelerate time to market; bolster data security; and reduce costs. By offering platform life cycle support, change-management services, and engineering consultation, Dell helps ensure that hardware-based appliances perform seamlessly in enterprise IT infrastructures.

For many IT organizations, the complexity involved in deploying critical data center capabilities and applications on shared servers can slow the rollout of new services. The trickle-down effect can impact the entire enterprise by prolonging product development cycles, complicating the delivery of customer support, compromising data security, and increasing data center operating costs.

These challenges are partly the result of the uncertainty that solutions providers face when they design products that must be compatible with a variety of potential data center environments. Makers of custom hardware must spend time acquiring expertise in designing and manufacturing hardware platforms, which can take their focus away from their core competency—for example, enterprise security or network forensics. Similarly, software-only solution vendors face the complexity of developing, testing, and supporting their products on several hardware platforms and an unpredictable range of potential software/software and software/hardware conflicts.

Taking control of critical processes

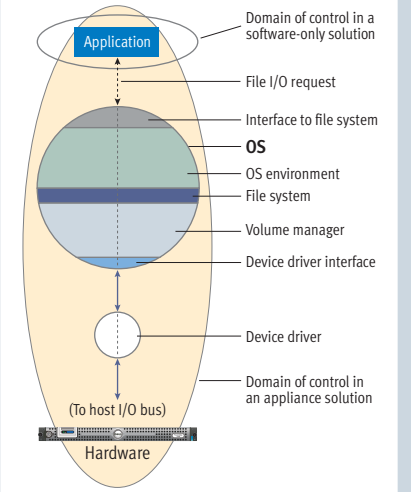
Enterprise IT organizations always have the option of deploying shrink-wrapped software that has been tested and qualified for a variety of hardware and software platforms. This model can work well for noncritical applications or those that are simple, robust, and flexible enough to withstand OS problems or conflicts with other applications. However, for complex applications that require a controlled or customized OS and hardware environment, it is often desirable to deploy software that has been "locked down" on one dedicated hardware platform.

To decide between a dedicated appliance and a software-only solution, IT organizations must consider the suitability of a dedicated hardware platform for its intended task. Common areas in which appliances can enhance the efficiency, manageability, and flexibility of data center operations include the following:

- Web security (such as firewalls, antivirus, and intrusion detection)
- Messaging security (such as antispam and content security)
- Network management
- Intranet search
- Enterprise applications (such as business-to-business middleware, database key management, and compliance)
- Network-critical applications that require an isolated and controlled OS and hardware environment to avoid bandwidth contention and conflicts (such as wide area network [WAN] acceleration or any application that is "in line" to the network traffic flow)

Helping reduce conflicts and boost security

A simple file request in a software application can help illustrate the potential benefits of an appliance model. Figure 1



Source: Building Storage Networks, by Marc Farley (McGraw-Hill/Osborne Media, 2000).

Figure 1. *The appliance model provides a controlled environment for data, helping eliminate the risk of hardware and software conflicts*

shows the typical data path from an application to the hardware. In this case, the application initiates the request through the file system interface. Beyond that point, the data flows through elements that are not directly controlled by the application software provider. The OS, device drivers, and underlying hardware are all controlled by the enterprise IT department and may cause conflicts with the software.

In the case of an appliance, each of these elements is controlled and locked down by the appliance vendor. In this way, appliances help eliminate the risk of conflicts with hardware or other software that are inherent in software-only solutions. Because appliances offer a known and controlled environment on which the application runs throughout its life cycle, they help provide enterprises with simple, consistent deployments and service on a known valid configuration. As a result, IT administrators can tune applications to a specific hardware set rather than to the lowest common denominator on a shared server. In addition, simplified support and troubleshooting on a single dedicated hardware platform can allow IT departments to resolve problems quickly—helping maintain high availability levels and reduce expensive downtime.

Besides avoiding potential software conflicts, appliances enhance data center performance. Because they do not compete for cycles with other software on a shared server, appliances can optimize performance for critical applications. They help bolster security because the applications they support are in an isolated operating environment that is not exposed to other applications or users. Appliances can also incorporate critical security technologies that mitigate the risk of data theft and help protect vital intellectual property. In addition, appliances help increase efficiency, simplify management, and facilitate growth by providing one location to manage, store, and monitor policy settings and audit information.

Making the "build or buy" decision

Before developing a dedicated appliance, solutions providers must evaluate the technical, financial, and life cycle considerations of a custom device that they would design and build or have an original design manufacturer build for them, compared with an appliance based on a customized off-the-shelf commercial hardware platform. Many large enterprises initially consider creating their own appliances because off-the-shelf products lack one or two critical features. However, thanks to advances in the performance of standards-based general-purpose hardware, dedicated appliances based on off-the-shelf servers have become an attractive alternative to custom devices designed from scratch (see Figure 2). Today, general-purpose server platforms such as those offered by Dell can provide ample computing power and can be adapted to many different types of applications with just a little customization.

As depicted in Figure 3, switches and routers occupy the lower layers of the Open System Interconnection (OSI) model. Routers and switches have been sold as appliances almost since their introduction—but over the past several years, vendors have aggressively moved applications in the upper layers of the OSI model to appliances as well. Widespread adoption of Web technologies and open standards such as the IP suite, XML, and Session Initiation Protocol (SIP) has accelerated this trend.

Addressing the platform life cycle

Dedicated appliances based on off-the-shelf commercial hardware platforms offer IT organizations a long and predictable life cycle—one that is not tied to a particular original design manufacturer's development cycles. As software evolves and new releases come out, custom hardware may not be able to keep up with the requisite processor, memory, or I/O requirements. Dell addresses such life cycle concerns by supplying long-term visibility and predictability of hardware features and costs, and a change-management support structure for original equipment manufacturers (OEMs) that offer Dell™ hardware–based appliances. In addition, the Dell supply chain

Software-only solution running on a shared server	Dedicated appliance	
	Software running on a custom server	Software running on an off-the-shelf commercial server
IT benefits • Does not require deployment or management of separate hardware • Does not require fulfillment or return material authorization (RMA) logistics for dedicated hardware	**IT benefits** • Allows comprehensive control over hardware, firmware, and software • Helps increase reliability by avoiding hardware and software conflicts	**IT benefits** • Allows comprehensive control over hardware, firmware, and software • Helps increase reliability by avoiding hardware and software conflicts • Can easily keep up with evolving hardware and infrastructure requirements • Enhances availability by avoiding complications that may arise when deploying, managing, and supporting specialized hardware from an original design manufacturer
Disadvantages • May cause conflicts with hardware or with other software • Competes with other software for processor cycles, reducing performance compared with dedicated appliances	**Disadvantages** • May not be able to keep up with evolving hardware and infrastructure requirements • May result in delays in resolving hardware problems when supported by low-volume original design manufacturers	

Figure 2. *Dedicated appliances based on off-the-shelf commercial hardware offer significant advantages over both software-only solutions and custom hardware*

facilitates fast, flexible growth by enabling high-volume server deliveries at exceptional price points, as well as an established framework for reliability testing to help ensure top performance of integrated applications.

By extending support to appliance OEMs as well as enterprise IT departments, Dell helps ensure that Dell hardware–based appliances will continue to be supported as the enterprise IT infrastructure evolves. Dell supports its appliance OEMs with offerings designed to enhance overall IT efficiency, manageability, and flexibility in enterprise data centers:

- **Product development:** Regulatory support, custom bill of material maintenance, custom factory integration management, and custom services
- **Logistics:** Program governance services, hardware evaluation support, order management, Dell Merge Center support, issue resolution and escalation, reporting, and logistics support
- **Engineering:** Long-range component-level road maps and block releases, access to Dell engineering and technical escalation processes, and vertical specialization

Assessing technical advantages

Dedicated appliances based on industry-standard Dell server platforms enable outstanding performance, flexibility, and life cycle management—and benefit from in-house expertise honed through Dell efforts to simplify IT throughout its own global operations. In addition, customer-driven design enhancements, partnerships with industry leaders, and an intense focus on operational improvements help Dell create and refine platforms designed to be both economical and highly reliable.

APPLIANCES IN ACTION

IronPort: Partners in stopping crime

When a solutions provider chooses a hardware platform for its products, it makes the decision with the end user—the enterprise IT organization—in mind. "Customers are not shy about asking what is under the hood," says Steve Shray, senior vice president of engineering and operations at IronPort Systems. "We are proud to say it's Dell. It is not uncommon for customers to wonder how reliable the platform is. Once they know it's Dell, that removes any doubt."

When the IronPort founders first put together the company strategy, they decided to stay out of the business of hardware design or hardware life cycle management. "Our core competency is on the software side," Shray acknowledges. "It became clear early on that Dell had the most compelling server platform solution for us. Dell was able to offer us the three key elements we needed: a low-risk inventory management agreement, a price advantage for a well-trusted and tested platform, and a service model that could not be touched by competitors. Combined, these three components cinched the deal and made it an easy decision to name Dell as the sole-source hardware provider for the full portfolio of IronPort e-mail security appliances."

Google: In search mode

Google chose Dell for its appliances that allow companies to index, organize, and retrieve information inside the corporate firewall. "We wanted to benefit from a company whose core competencies would free Google to stay focused on its own strengths," explains Aidymar Bigio, manufacturing and operations manager at Google. "With excellent engineering and service, as well as proficiency in all facets of supply chains, the Dell OEM division provided us with no-excuses dependability, which extends to every one of our customers." The Dell OEM Industry Solutions Group provides extensive customization to Dell PowerEdge 2950 servers running the Linux® OS to form the foundation for the Google Search Appliance.

"Dell is enabling us to go to market quickly while scaling faster—in more locations—and all with a lean staff."

—Aidymar Bigio
Manufacturing and operations manager at Google
June 2007

"When any consumer has a bad experience with any product, the repercussions are significant," says Bigio. Dell enhances the customer experience by helping Google exceed service-level agreements. In addition, Dell helps improve supply chain management and customer fulfillment capabilities. "Dell is enabling us to go to market quickly while scaling faster—in more locations—and all with a lean staff," says Bigio.

Ingrian: Securing success

For an updated generation of its DataSecure appliances, Ingrian Networks decided on custom-branded Dell PowerEdge 1950 and PowerEdge 2950 servers as a cost-effective hardware platform for encrypting critical data in applications and databases. "We found the Dell PowerEdge servers to be very sturdy, well designed, and able to deliver the performance and reliability needed to handle complex encryption algorithms for mission-critical data. Those are the things that make a difference for our customers," says Brian Sheredy, director of business operations at Ingrian. Streamlined implementation, scalability, and flexible multi-tier integration are also key criteria, helping organizations maintain compliance with legislative and policy mandates for security.

The basic elements of general-purpose servers—processors, memory, and I/O—have become quite robust over time regardless of vendor. However, ninth-generation Dell PowerEdge servers offer several key features that can help simplify deployment, support, and operations for enterprise data centers. PowerEdge 1950, PowerEdge 2950, and PowerEdge 2970 servers and PowerVault MD1000 storage systems offer comprehensive, customizable platforms for appliances solutions, which can help keep integration costs low.

For example, Dell products offer a factory-configurable choice of PCI Extended (PCI-X) or PCI Express (PCIe) slots. Programmable LCDs can provide fault alerts, while a cable-less motherboard that directly connects to hard drives can help improve reliability and reduce heat inside the servers. Dell-backed service and support for appliance platforms can help minimize the impact of product transitions, enhance reliability, minimize maintenance costs, maintain consistent product performance, and preserve product stability.

Integrated Dell solutions are designed to use open standards, which helps keep costs low by taking advantage of broad industry adoption. Standards can help drive cost out of the design and build phases, allowing OEMs of Dell hardware–based appliances to pass savings on to enterprise IT organizations. In addition, open standards help ensure high reliability, compatibility, and integration of the solution across the entire IT environment.

Conducting extensive reliability tests

The Dell Reliability Test Facility strives to support high levels of reliability and availability for appliance platforms. In a dedicated 130,000-square-foot facility with state-of-the-art equipment designed to test shocks and vibrations, packaging, safety, and thermal stress and to perform failure analysis, Dell tests its models and methodologies based on real-world usage scenarios. This practice helps ensure top reliability, compatibility, and integration reaching from the desktop to the data center.

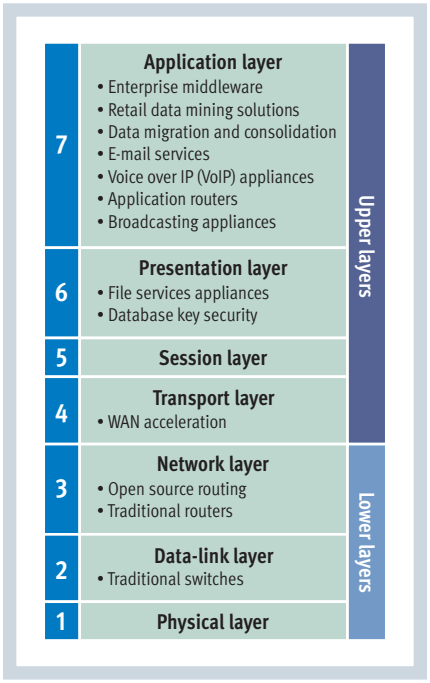


Figure 3. *Appliances can be used in both the lower and upper layers of the OSI model*

Moreover, the Dell open systems design philosophy and ongoing support to its appliance OEMs help ensure smooth operations and simplified IT management. Dell takes a holistic approach to hardware and compatibilities within data center environments and provides OEM-specific services and support to over 40 industry verticals—giving appliance OEMs a critical support structure that is transparent to enterprise users. For details on how three companies have taken advantage of Dell hardware to create dedicated appliances, see the "Appliances in action" sidebar in this article.

Simplifying the IT environment

Because millions of Dell systems are in use worldwide, Dell is equipped to recognize and resolve hardware issues much sooner than providers such as low-volume original design manufacturers. This real-world experience can benefit enterprise IT departments by helping simplify installations, reduce complexity, and enhance serviceability. In addition, Dell offers worldwide on-site support.

Dell can also fulfill demand quickly, enabling IT organizations to receive appliances and deploy new capabilities rapidly after placing an order. Through its outstanding supply chain execution and customization capabilities, Dell is well equipped to meet demand variability and deliver on time directly to enterprises around the world.

Running a data center is never an easy job—but dedicated appliances built on standards-based, off-the-shelf commercial platforms can play a pivotal role in simplifying IT. Through top-notch product quality, support for Dell appliance OEMs, and worldwide support, Dell helps enterprises deploy critical data center capabilities quickly and easily—simplifying support while bolstering data security.

Jaleel Mohammed *is a senior systems consultant in the Dell OEM Industry Solutions Group. He has 12 years of experience in the embedded systems industry in Silicon Valley. Jaleel has a degree in Computer Science and Mathematics from the University of Wisconsin at Madison.*

David Pennington *is a system consultant manager in the Dell OEM Industry Solutions Group. David has been at Dell for more than 10 years, previously working in product operations and marketing. He attended the University of Texas.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell OEM appliance solutions:
DELL.COM/downloads/us/biz/Dell_OEM_Appliance_Slicksheet.pdf

Dell OEM Industry Solutions Group:
DELL.COM/OEM

Dell Infrastructure Consulting Services:
DELL.COM/ICS

HIGH-STAKES TECHNOLOGY: HOW DELL PLATFORMS KEEP CASINO VIDEO GAMES FLUSH

BY PATTY YOUNG
FRANKLIN FLINT
DAVID HAVENER

As video-based gaming technology surges in popularity, Dell™ server and storage platforms are behind the scenes at many casinos, helping reduce management complexity while enhancing the reliability, serviceability, and scalability of highly customized gaming environments.

Gaming is big business in many countries around the world, and sophisticated technology platforms keep the video-based gaming business running. For example, according to the American Gaming Association, gross gambling revenues in the United States reached US$84.65 billion in 2005[1]—and according to PricewaterhouseCoopers, the industry generated more than US$1.3 trillion in 2005.[2] More than a quarter of the U.S. adult population visited a casino in 2006.[3] And gaming is not limited to large, well-known gambling venues such as Las Vegas, Atlantic City, and Monaco; numerous Native American casinos have become popular across the United States as well.

Behind the scenes, video gaming devices rely on high-performance, high-availability hardware and software platforms. Not only are there many mechanical and computer components within the gaming cabinets themselves, but multiple systems run constantly in the background to gather tracking data from each of these machines to help casinos keep tabs on their gaming floors. This data—which may include information on wins, payouts, and number of plays—must be gathered and archived for reporting and audits.

As video gaming technologies have become widely available, many in the gaming industry are attempting to embrace standards for communication between video gaming devices and the back-end systems supporting them, such as servers, databases, and data mining applications. These new video gaming standards are expected to borrow from existing standards such as XML, Simple Object Access Protocol (SOAP), and Web services. Some gaming technology providers are now cooperating in this endeavor, with the goal of making their video gaming software extensible for the future. Through software standards such as Games to Servers (G2S), video gaming technology vendors plan to reduce the complexity of adding new game features in the future.

Minimizing management complexity

The prospect of running additional communications cables is not an attractive one for most casinos, who prefer to leverage existing Ethernet connections through network designs with a single wire to the floor for each slot machine. Dell hardware helps casinos achieve these designs by enabling bidirectional communications to the floor. In addition, Dell is working to provide real-time monitoring of slot machine components to help optimize uptime.

Besides casinos, many video gaming technology vendors utilize Dell hardware and services to support their software applications. Gaming software performance relies heavily on the hardware platform running it. IT complexity can lead to planned or unplanned downtime, often resulting in lost revenue, compromised tracking data, and disgruntled customers.

[1] "Gaming Revenue: Current-Year Data," by the American Gaming Association, October 2006, www.americangaming.org/Industry/factsheets/statistics_detail.cfv?id=7.

[2] "Global Entertainment and Media Outlook: 2007-2011-Casino and Other Regulated Gaming," by PricewaterhouseCoopers, June 21, 2007, www.marketresearch.com/product/display.asp?productid=1521264&xs=r.

[3] "2007 State of the States: The AGA Survey of Casino Entertainment," by the American Gaming Association, www.americangaming.org/assets/files/aga_2007_sos.pdf.

It is critical to casinos and video gaming technology vendors that their games work properly while the casino is open for business—which is typically 24 hours a day. Dell helps address this requirement by including common design elements across many of its Dell PowerEdge™ server models, helping minimize change-management costs and enhance productivity through intuitive systems management interfaces and tools.

Integrated gaming solutions based on industry-standard Dell server and storage systems are also designed for ease of maintenance and cost-effective scalability. Cost-efficient Dell hardware platforms are designed to provide the performance, uptime, and resilience required to help simplify IT management in video gaming environments. In addition, the Dell OEM Industry Solutions Group can work with video gaming vendors to configure Dell hardware to optimize the performance of custom software. Well-matched hardware and software solutions help ensure that the casino network topology remains transparent to gaming customers.

Creating a resilient IT infrastructure

High availability is critical for video gaming devices because in a casino, time is money—literally. When access to critical functions is compromised, the result can be lost revenue, lost productivity, customer dissatisfaction—and, in extreme cases, the possibility of data loss. To help avoid these pitfalls, the Dell PowerEdge servers and Dell PowerVault™ storage used in typical high-end video gaming deployments have been designed with a broad array of redundant features to maximize system availability.

For example, Dell helps provide gaming configurations that include Marathon everRun HA high-availability solutions, which are built using industry-standard off-the-shelf components to support three levels of availability. Entry-level non-clustered servers, mid-range clusters, and high-end clusters have all been designed to remove single points of failure. Advancing from one availability level to the next enables casinos to recover from an increasing number of potential component failures.

Designed to deliver maximum computing power with minimum complexity, Dell PowerEdge 2950 servers deliver outstanding performance combined with internal expandability in a rack-dense 2U form factor. PowerEdge 1950 servers provide the same advantages in a 1U form factor and share a common system image with PowerEdge 2950 servers, which helps lower management costs by simplifying software component updates. In addition, PowerEdge 2900 servers act as image compatibility platforms, with up to two quad-core Intel® Xeon® 5300 series processors delivering performance gains of up to 63 percent over the previous generation of servers.[4] As a result, PowerEdge 2900 systems can provide additional processing power when an incremental increase is required.

Dell storage systems—particularly PowerVault MD1000 and PowerVault MD3000 storage systems—play an important role in supporting video gaming technology as well. The modular design of the PowerVault MD1000 disk storage expansion enclosure allows video gaming technology vendors to handle growing data volumes seamlessly. To help keep management simple, the PowerVault MD1000 supports up to 15 Serial Attached SCSI (SAS) and Serial ATA (SATA) drives. One PowerEdge server equipped with a single PowerEdge Expandable RAID Controller (PERC) 5/E can daisy-chain up to six PowerVault MD1000 enclosures for a total of 90 hard drives.

The PowerVault MD3000 is a high-performance modular disk storage array that supports up to 15 hot-pluggable SAS drives for applications running on one, two, or a pair of clustered PowerEdge servers. Designed with redundant components and connections to help mitigate risks of downtime, the PowerVault MD3000 storage array can withstand hardware failures, broken cable connections, and similar problems.

Figure 1 illustrates a typical high-end video gaming deployment from Dell integrating these components. It consists of two Marathon everRun HA clusters of PowerEdge 2950 servers and PowerVault MD3000 storage, all in a rack-dense

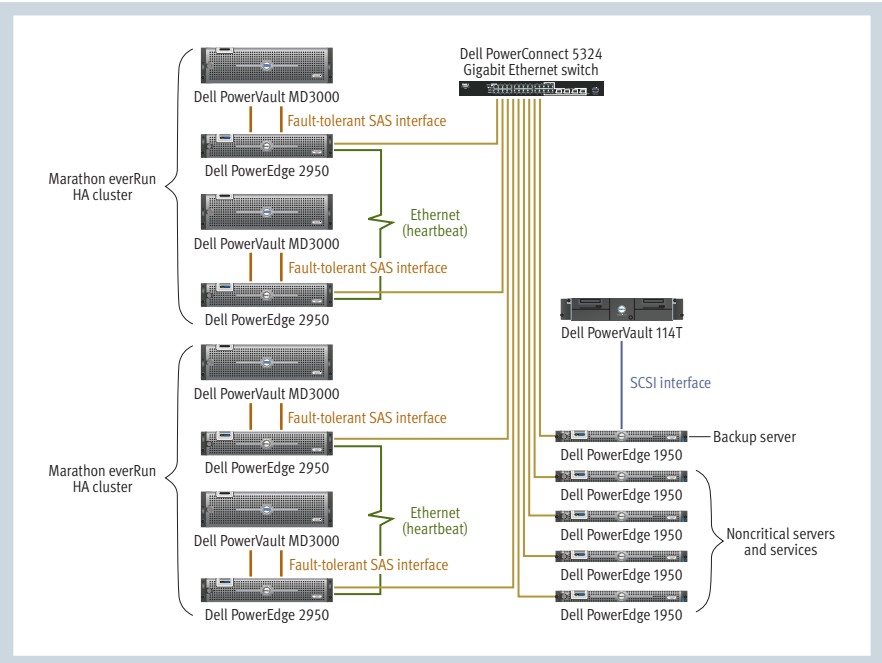


Figure 1. *A resilient IT infrastructure helps keep high-end video gaming deployments running 24/7*

[4] Based on the SPECint_rate2000 benchmark test performed by Dell Labs in May and October 2006 on a Dell PowerEdge 2950 with two quad-core Intel Xeon X5355 processors at 2.66 GHz; 8 GB of 667 MHz fully buffered dual in-line memory modules; one 80 GB, 7,200 rpm SATA hard drive; and the Microsoft Windows Server® 2003 Enterprise x64 Edition OS, as compared with a similarly configured PowerEdge 2950 with two dual-core Intel Xeon 5160 processors at 3.0 GHz. Actual performance will vary based on configuration, usage, and manufacturing variability.

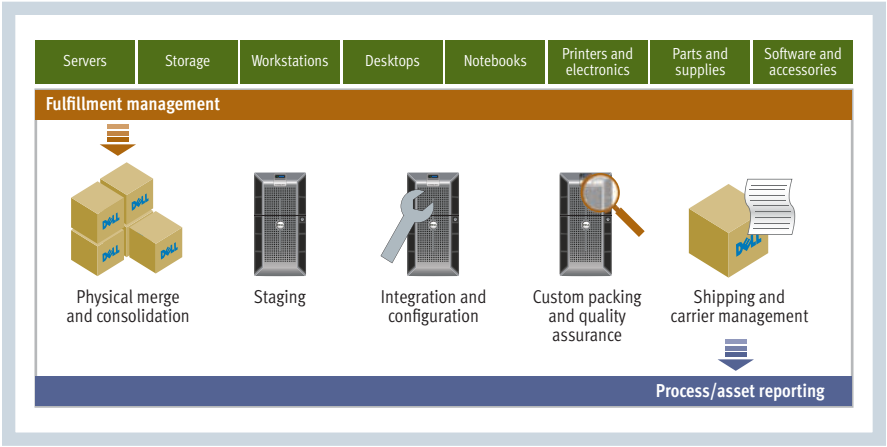


Figure 2. *Dell Merge Centers combine configuration, custom packing, shipping, and carrier management to streamline the deployment process and help lower costs for video gaming solutions*

2U cluster configuration offering high availability as well as the option of dual-core or quad-core processors. These clusters connect through a Dell PowerConnect™ 5324 Gigabit Ethernet switch to additional on-site PowerEdge 1950 standby servers for failover in the event of a hardware failure and other noncritical services, as well as a PowerVault 114T tape rack enclosure for backups. This type of deployment may also incorporate the Microsoft® SQL Server™ database platform, tape backup software, remote management tools such as Symantec pcAnywhere software, and other applications, as well as rack hardware and peripherals. These Dell systems can scale for various environments—for example, one organization simply adds a PowerVault MD1000 storage system for each increment of 500–1,000 slot machines, expanding its SQL Server cluster and performance with additional spindles.

Reducing on-site setup time

The Dell OEM Industry Solutions Group complements Dell hardware offerings by providing "rack-and-stack" services through its Merge Centers, which are designed to ease deployment and reduce costs (see Figure 2). Dell Merge Centers offer a logistical assembly point where video gaming hardware can be received, unboxed, racked, cabled, mounted, and tested before the integrated solution is shipped. By using Dell Merge Center capabilities, gaming vendors can get a staging and assembly point where much of a large-scale deployment can be built away from the casino floor, thereby helping reduce on-site setup time, limit the amount of shipping container waste, and minimize business disruptions. A single point of contact also helps streamline the deployment. In addition, Dell Services offers a range of options to fit specific organizational needs, including Silver-, Gold-, and Platinum-level support.

After identifying the needs of the video gaming technology vendor, Dell performs custom configurations during the initial system build. The Dell "one-touch" process of custom integration enables systems to be built and shipped only once through factory or multiple-vendor channels. After configuration, Dell products are delivered directly from a Dell International Organization for Standardization (ISO) 9001–certified factory.

Standard or custom configurations of standard, custom, or proprietary software can also be installed in the Dell factory environment based on an organization's requirements and schedule. In this way, video gaming technology vendors can help ensure that their systems have the same version of the same software.

Standardizing on a tiered approach for growth

As the video-based gaming industry continues to expand, technology providers require reliable, cost-effective hardware platforms capable of supporting 24/7 operations. By standardizing on Dell hardware platforms, video gaming vendors can leverage the outstanding performance, high availability, and management simplicity of tested Dell server and storage systems while maintaining their own branding and support with casinos and other gaming establishments. In addition, a tiered approach allows organizations to implement suitable hardware and services to meet specific requirements while enabling a pay-as-you-grow strategy to help control costs.

Patty Young *is a senior systems consultant in the Dell OEM Industry Solutions Group specializing in server and storage products. She has a degree from North Carolina State University and has been working in the computer industry for 22 years.*

Franklin Flint *is a senior systems consultant in the Advanced Systems Group supporting the Dell OEM Industry Solutions Group. He has been at Dell for 12 years, focusing on original equipment manufacturer sales for the past 6 years.*

David Havener *is a systems consultant in the Dell OEM Industry Solutions Group specializing in servers and storage. He has a degree from the University of Minnesota and has been with Dell for more than 10 years.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell Merge Centers:
DELL.COM/content/topics/global.aspx/services/adi/dps_hpcc#Merge_Center

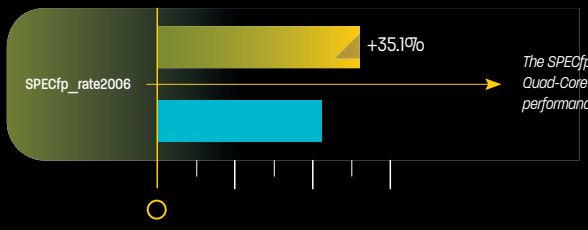
Dell OEM Industry Solutions Group:
DELL.COM/ISG

Dell Services:
DELL.COM/Services

Leading the industry again, this time in quad-core.

The new Quad-Core AMD Opteron™ Processor delivers much more than just four cores.

1 Industry-leading performance-per-watt. For industry-leading savings. 2 Optimal virtualization technology. For getting the most out of every server. 3 Consistent architecture, consistent power and thermal envelopes. For superior investment protection. 4 Native quad-core with Direct Connect Architecture. For a whole new level of performance. But don't just take our word for it. Feel free to explore. Visit www.amd.com/quadcore



The SPECfp_rate 2006 peak score shows the Quad-Core AMD Opteron™ Processor enjoys a full 35.1% performance advantage over the Xeon 5345.

AMD
Smarter Choice

Building a Fast, Flexible Platform for Virtualization with AMD Opteron™ Processor–Based Dell PowerEdge Servers

Dell™ PowerEdge™ servers with Quad-Core AMD Opteron™ processors help drive down total cost of ownership while accelerating business agility.

Dell PowerEdge 2970 and PowerEdge 6950 servers featuring Quad-Core AMD Opteron™ processors

Virtualization gained momentum in the data center by enabling organizations to consolidate servers, reduce hardware and administration costs, and optimize performance. But today's exacting service-level requirements call out for more: a virtual IT infrastructure that enables rock-solid business continuity within a highly scalable, highly available framework. Quad-core processor technology helps meet these objectives by significantly enhancing performance and energy efficiency in cost-effective, industry-standard servers.

Dell has partnered with AMD and major virtualization management suppliers to help data center managers achieve the full benefits of virtualization. Dell PowerEdge servers configured with Quad-Core AMD Opteron™ processors are designed for fast business response and flexible change management—providing the traction to support high availability, disaster recovery, and data protection while helping reduce total cost of ownership (TCO).

Designed for acceleration

Ready, set, go. With fully certified virtualization solutions, Dell is making it easy to capitalize on virtualization opportunities across the enterprise using PowerEdge 2970 and PowerEdge 6950 servers. Featuring Quad-Core AMD Opteron processors, these servers are well equipped to address data center challenges associated with virtualization, such as cost-effective memory scaling, storage utilization and complexity, dynamic resource provisioning, zero-downtime maintenance, and incremental, demand-based I/O bandwidth increases:

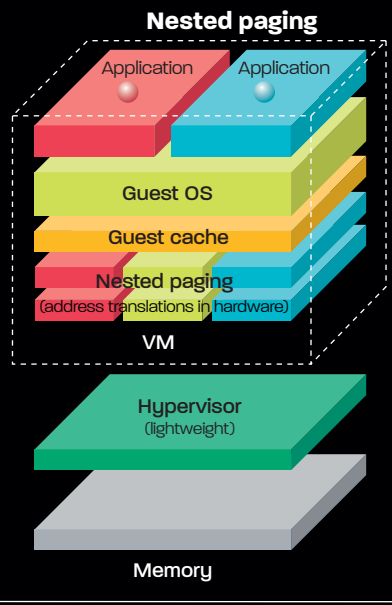
- **Dell PowerEdge 2970** servers appeal to savvy but cost-conscious data center administrators, featuring Dell's energy-efficient server design together with AMD PowerNow!™ technology to help optimize processor system power utilization. This 2U server is well equipped to advance virtualization for mainstream business applications while helping reduce energy costs.
- **Dell PowerEdge 6950** servers are designed to harness maximum performance with minimum power consumption and are well suited for virtualizing back-end database environments and custom applications. This four-socket server uses high-performing AMD Opteron processors to enable highly efficient data transmission, availability, and reliability for mission-critical enterprise applications.

In addition, Dell is planning to release a new server built from the ground up for virtualization. Incorporating third-generation Quad-Core AMD Opteron processors and an embedded hypervisor, these servers enable IT departments to deploy virtualized systems quickly and easily, right out of the box. And by combining high-efficiency processors, power supplies, and fan technology with the potential for high consolidation ratios, these servers can help significantly reduce both energy costs and TCO.

AMD
Smarter Choice

AMD Virtualization™ technology enhancements accelerate performance

- Each guest OS manages its own environment
- Guest cache reduces nested paging translations
- Memory lookups done in hardware are typically faster than those done in software
- The hypervisor no longer maintains shadow copies of page tables



Engineered for performance

To meet the relentless demands of always-open-for-business operations, organizations require rugged, tested solutions. That is why these Dell PowerEdge servers use nimble Quad-Core AMD Opteron processors to help enhance security, performance, and efficiency in virtualized IT environments.

AMD's multi-core technology is designed to host multiple virtual machines (VMs) per processor while maintaining consistent thermal and electrical envelopes, helping boost performance to leverage existing investments. In addition, Quad-Core AMD Opteron processors offer several distinct advantages for optimizing virtualization, including Direct Connect Architecture, efficient double data rate 2 (DDR2) memory and dedicated level 2 (L2) cache, Dual Dynamic Power Management™ technology, and Rapid Virtualization Indexing. With the introduction of the PowerEdge 2970, Dell was the first tier 1 server vendor to bring AMD Dual Dynamic Power Management to market.

Direct Connect Architecture. Using an integrated memory controller and HyperTransport™ technology for fast CPU-to-memory and CPU-to-CPU communication, Direct Connect Architecture helps eliminate bottlenecks, reduce latency, and shorten world switch times—that is, the time it takes to switch between VMs. A switch typically commands up to 2,000 processor cycles, but new instructions built into Quad-Core AMD Opteron processors help shorten cycles significantly. Because virtualization is memory intensive, fast access to memory has become a key enabler for increased virtualization performance. The integrated memory controller in Direct Connect Architecture enables faster memory access than traditional frontside bus–based architectures, which are not designed with an integrated memory controller.

Efficient DDR2 memory and dedicated L2 cache. By helping speed storage performance for working data, DDR2 memory is integral to AMD Opteron processors' lightning-fast response in virtualized storage environments. This translates into high peak throughput performance when it is needed most—from an architecture that is designed to reduce power consumption and speed communication times compared with other memory technology such as fully buffered dual in-line memory modules (FBDIMMs).

L2 cache memory is designed to act as a staging bunker to supply data and instructions at higher speeds than main memory. Because L2 cache typically shares memory, when multiple VMs try to access the same memory pool, the effectiveness of retrieval is often diminished or deadlocked; this is sometimes referred to as cache pollution. In contrast, AMD Opteron processors use dedicated L2 cache, designed to improve virtualization performance by providing a unique workspace for each processor core—thereby avoiding the lag times of shared cache memory. The addition of level 3 (L3) cache on Quad-Core AMD Opteron processors provides supplemental overflow workspace for cores processing extra-large data sets.

Dual Dynamic Power Management™ technology. A new feature of Quad-Core AMD Opteron processors, and designed from the ground up into the Dell PowerEdge 2970 server, Dual Dynamic Power Management helps optimize energy efficiency to lower TCO. This is accomplished using a "split plane" voltage system designed to reduce memory latency for increased application performance in quad-core processors compared with dual-core processors. In addition, Dual Dynamic Power Management coordinates with AMD PowerNow! technology to enable further reductions in power and cooling requirements and advance energy efficiency. Moving forward, Dell plans to design AMD Opteron processor–based PowerEdge servers to support Dual Dynamic Power Management.

Rapid Virtualization Indexing. Another AMD technology feature that Dell servers are expected to natively support in the near future, Rapid Virtualization Indexing includes nested paging and tagged TLBs to help increase performance, efficiency, and flexibility for many virtual workloads. Nested page tables are used in the translation of virtual addresses to physical addresses. Working with independent software vendor (ISV) virtualization partners, AMD has designed this approach to minimize the complexity of memory management in virtualized environments. AMD accomplishes this through a two-step process: first, the VM address is mapped to a guest physical address, which is then mapped to a physical server address. Direct memory management enhances virtualization performance by reducing the intervention required from the virtualization management software, or *hypervisor.*

Built for business agility

Dell PowerEdge servers and AMD Opteron processors help increase overall business agility by capitalizing on a virtualization strategy designed to reduce complexity, improve performance, and sustain value across the enterprise. Furthermore, energy-efficient AMD Opteron processor–based PowerEdge servers are designed from the ground up for expert handling of the virtual IT infrastructure—working with existing investments to help cost-effectively boost performance without reengineering the data center.

Find out more about the advantages of AMD Opteron™ processor–based Dell PowerEdge servers now!

AMD Opteron™ processors:
www.amd.com/opteron

Dell PowerEdge servers:
DELL.COM/Enterprise

SYMANTEC ENDPOINT PROTECTION: A UNIFIED, PROACTIVE APPROACH TO ENTERPRISE SECURITY

BY LAUREN DUDA

Enterprises must constantly guard against stealthy, targeted, financially motivated attacks that can exploit vulnerabilities in endpoint devices to allow data theft and other damage. Through a seamless, multilayered approach, Symantec® Endpoint Protection helps provide advanced threat prevention and simplified, holistic endpoint protection across notebooks, desktops, and servers.

Related Categories:

Antivirus

Firewalls

Security

Symantec

Threat management

Visit ***DELL.COM/PowerSolutions*** *for the complete category index.*

IT threats have changed dramatically over the past few years, as attacks with the simple goal of making headline news have given way to a relentless wave of financially motivated attacks against unsuspecting enterprises. With this goal in mind, professional hackers continuously develop sophisticated new tactics to escape detection and discover new entry points that allow them unauthorized and ongoing access to an organization's systems and valuable information.

Antivirus, anti-spyware, and other signature-based protection measures may have been sufficient to protect an organization's vital resources a few years ago, but not today. Although these primarily reactive methodologies can still play a vital role in protecting notebooks, desktops, and servers, they are only part of the solution. To help protect these endpoints against a comprehensive array of sophisticated threats—including unknown threats and zero-day attacks—organizations must augment their traditional security approach with proactive endpoint security.

However, for many organizations, adding proactive endpoint security can translate into installing discrete antivirus, anti-spyware, desktop firewall, intrusion prevention, and other types of software from multiple vendors. Deploying this software individually on each endpoint not only is time-consuming, but also increases IT costs and complexity, and typically requires providing management, training, and support for each application. In addition, these different applications can actually work against one another to create security gaps, and may require high resource consumption that can impede system performance.

To help eliminate security gaps, increase endpoint system performance, and reduce the costs and complexities associated with deploying and managing multiple endpoint protection solutions, Symantec Endpoint Protection consolidates multiple endpoint protection technologies into a single integrated agent that administrators can control from a unified management console. Symantec Endpoint Protection combines industry-leading antivirus and anti-spyware signature-based protection with firewall, device control, and proactive intrusion prevention software, and employs state-of-the-art threat prevention to help protect against known and unknown malware, including viruses, worms, Trojan horses, spyware, and adware. It even helps protect against sophisticated attacks that evade traditional security measures such as rootkits, zero-day attacks, and mutating spyware.

Advanced antivirus and anti-spyware technology

Although typically inadequate for protecting against unknown threats and zero-day attacks, antivirus and anti-spyware software is still essential to endpoint security. This software typically uses traditional scan-based technologies to identify viruses, worms, Trojan horses, spyware, and other

> *"Symantec Endpoint Protection consolidates multiple endpoint protection technologies into a single integrated agent that administrators can control from a unified management console."*

malware on an endpoint device—searching the system for files that match the characteristics, or threat signatures, of a known threat. Once these technologies detect a threat, they remediate it, typically by deleting or quarantining it. For many years, this type of software has been effective in helping protect endpoints against known threats.

With the IT industry's increased focus on endpoint security, a variety of antivirus and anti-spyware products have become available. Although many of these first- and second-generation solutions provide some protection, they can fall short of comprehensive protection—only working on one OS, for example, or lacking the ability to interoperate with other essential endpoint security elements such as personal firewalls, intrusion prevention systems (IPSs), and device control.

Antivirus and anti-spyware software may also be unable to detect polymorphic threats or detect and remove rootkits. For example, in a February 2007 study conducted by AV-Comparatives, of 15 antivirus products tested with 12 highly polymorphic viruses, only Symantec and one other vendor received a score of 100 percent for all polymorphic viruses tested.[1] According to the report, these tests help determine the flexibility of an antivirus scan engine as well as how good it is at detecting complex viruses.

Rootkits—stealth applications or scripts that hackers can use to gain undetectable administrator-level access to a system—are widely available on the Internet, giving inexperienced hackers easy access to these tools without having to understand how they work. Rootkits are often used to collect confidential information such as user IDs, account numbers, and passwords. Detecting and removing rootkits typically requires thoroughly analyzing and repairing the OS, something many antivirus solutions may be unable to do. To this end, Symantec Endpoint Protection 11.0 is designed to provide a deeper inspection into the file system than many other solutions, enabling the analysis and repair processes necessary to remove even highly difficult rootkit attacks.

State-of-the-art network threat protection

Network threat protection is critical to defending endpoints against blended threats and inhibiting outbreaks. To be effective, it must encompass more than just a firewall: it should combine state-of-the-art protection technologies, including intrusion prevention and sophisticated capabilities for controlling network communications.

In the past, security experts have debated whether organizations should place firewalls only on the perimeters of their networks, or on individual desktops as well. As threats have multiplied and mobile workforces have extended the perimeters of organizations' network infrastructures, however, endpoints have become a primary target for attacks: an exploit may first infect a single notebook outside the network perimeter, then spread to other endpoints when the notebook connects to the internal network. Endpoint firewalls have become key not only to blocking internal network attacks from breaching endpoints connected to the network, but also to helping prevent these threats from leaving the initially infected endpoint.

The Symantec Endpoint Protection single-endpoint security agent incorporates a state-of-the-art firewall that includes the following key components:

- Rule-based firewall engine
- Predefined antivirus, anti-spyware, and personal firewall checks
- Firewall rule triggers based on applications, hosts, services, and time periods
- Comprehensive TCP/IP support, including User Datagram Protocol (UDP) and Internet Control Message Protocol (ICMP)
- An option to allow or block support of network protocols, including the Ethernet, Token Ring, Novell® IPX/SPX™, Apple AppleTalk, and NetBIOS Extended User Interface (NetBEUI) protocols
- The ability to block protocol drivers such as those for VMware® virtualization software and Windows Packet Capture Library (WinPcap)
- Adapter-specific rules
- Encrypted and clear-text network traffic inspection
- Packet and stream IPS blocking, custom IPS signature blocking, and generic exploit blocking for proactive threat protection
- Self-enforcement for network access control

Because intrusion prevention is vital to helping protect against vulnerability-based intrusions that utilize generic signatures, it must play a critical role in network threat protection. However, although traditional IPSs can detect a specific known exploit, they are typically inadequate against the barrage of exploit variants that exist today. According to the March 2007 *Symantec Internet Security Threat Report*, from July to December 2006 it took 47 days on average for an OS or application

[1] "Anti-virus Comparative No. 13: On-Demand Detection of Malicious Software," by AV-Comparatives, February 2007, www.av-comparatives.org/seiten/ergebnisse/report13.pdf.

Threat	Single GEB signature	Number of variants blocked
Blaster	MS RPC DCOM BO	814
W32.Mytob.IM@mm	MS_RPC_NETDDE_BO	426
Sasser	MS LSASS BO	394
W97M.Invert.B	RPC_NETAPI32_BO	250
W32.Gaobot.AAY	NetBIOS MS NO (TCP)	121
Welchia	MS IIS Webdav Exploit	55
W32.Zotob.A	MS Plug and Play BO	51
W32.Welchia.C	MS Locator Service BO	43

Figure 1. *Multiple exploit variants blocked through generic exploit blocking signatures*

provider to release a patch for a published vulnerability.[2] Attacks that exploit these vulnerabilities before a patch becomes available are often referred to as unseen or zero-day attacks. A few hours after the first vulnerability exploit is detected, IPS vendors typically can release a signature to protect against further attacks from the specific exploit.

These reactive measures create significant opportunities for sophisticated attackers. Considerable damage can be inflicted on an organization with the first wave of exploits before the release of an exploit signature. Even after the exploit signature is released, it may prove ineffective against polymorphic or self-mutating variants of that exploit. Furthermore, these reactive, exploit-based signatures cannot protect against unseen, unreported, or unknown threats, such as stealthy exploits targeted at specific organizations, which often go undetected. Combating these types of threats requires proactive measures through a vulnerability-based IPS.

While an exploit-based signature detects only a specific exploit, a vulnerability-based signature operates at a higher level—it can detect not only a specific exploit for a vulnerability, but potentially any exploit that attempts to attack that vulnerability. Symantec Endpoint Protection includes generic exploit blocking (GEB), a vulnerability-based IPS technology that uses generic signatures. When OS or application vendors announce new vulnerabilities, Symantec engineers study the characteristics of that vulnerability, then create and release a generic signature based on that study. The power of vulnerability-based IPSs derives from the fact that a single vulnerability definition can protect simultaneously against multiple types of threats (see Figure 1). Because these definitions are based on vulnerability characteristics and behavior, they can help protect organizations against a wide range of threats—even those that are not yet known or developed.

Vulnerability-based IPSs are also useful for protecting against exploits that target a specific organization. Targeted attacks are generally stealthy, because the attackers' goal is to steal confidential information and then erase themselves from the system without being discovered. Because organizations have no way of knowing about these targeted exploits until the damage is done, there is no way to create an exploit signature that could have prevented the attack. Vulnerability-based IPSs can detect and block the exploit by recognizing the high-level characteristics of the vulnerability that the targeted attack is attempting to exploit. The endpoint security agent in Symantec Endpoint Protection incorporates vulnerability-based protection at the network layer, helping prevent even unseen exploits and exploit variants from entering and infecting the endpoint—avoiding damage and the need for remediation.

Symantec Endpoint Protection also enables administrators to create custom rule-based intrusion prevention signatures tailored to the needs of their specific environment. They can create signatures that block a few specific actions or multiple complex actions. By helping eliminate the need to wait for an OS or application vendor to create patches for known vulnerabilities, Symantec Endpoint Protection provides administrators with comprehensive, proactive control over endpoint security.

Proactive threat protection

While signature-based file and network scanning technologies address key areas of protection, non-signature-based approaches are also necessary to address the growing number of unknown threats used in stealth attacks. These approaches are referred to generally as proactive threat protection technologies.

Symantec Endpoint Protection includes Proactive Threat Scan, which is based on heuristics technology that analyzes the behavior of system processes to help protect against the

"Targeted attacks are generally stealthy, because the attackers' goal is to steal confidential information and then erase themselves from the system without being discovered."

[2] *Symantec Internet Security Threat Report, Volume XI: Trends for July–December 06,* by Symantec, March 2007, eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_internet_security_threat_report_xi_03_2007.en-us.pdf.

multitude of variant and unseen threats that exploit known vulnerabilities. Its host intrusion prevention capabilities also enable organizations to protect themselves against unknown or zero-day threats.

Many host-based IPSs only examine what they consider to be "bad behavior" by applications. As a result, they can often falsely identify acceptable applications as threats and shut them down, causing productivity problems for users and help-desk nightmares for administrators. Proactive Threat Scan, however, scores both good and bad behavior, helping increase the accuracy of threat detection and reduce the number of false positives.

Symantec Endpoint Protection also incorporates device and application control capabilities that allow administrators to deny specific activities deemed high risk and to block specific actions based on user location. Device control allows administrators to determine which devices are allowed to attach to an endpoint—for example, it can lock down an endpoint to keep USB drives, CD burners, printers, and other USB devices from connecting to help prevent confidential data from being copied off of the system. Its ability to block device connections can also help prevent endpoints from being infected by viruses spread from these and other types of devices.

Application control allows administrators to control access to specific processes, files, and folders by users and other applications. It provides application analysis, process control, file and registry access control, and module and dynamic-link library (DLL) control. These advanced capabilities are useful for administrators who want to restrict certain activities deemed suspicious or high risk.

Integrated network access control

The Symantec Endpoint Protection agent integrates network access control, which administrators can easily enable by purchasing a Symantec Network Access Control license. After deploying Symantec Endpoint Protection, administrators do not need to deploy additional agent software on endpoint devices to implement network access control.

Convergent security and management

Effective client management is key to reducing costs and providing a quality, stable, and secure computing environment. Dell Client Manager™ software, powered by Altiris, can enhance security configurations to help make well-managed endpoints into secure endpoints as well. Incorporating security functionality into their management console helps organizations increase productivity while reducing endpoint management costs.

The Symantec Endpoint Protection Integration Component, which Altiris plans to include in all Dell Client Manager versions, facilitates migration to Symantec Endpoint Protection–based clients using remote delivery mechanisms. This component provides detailed reporting, broad dashboard deployment views, multicasting technology, advanced client discovery, and the ability to scale for both LAN-connected and remote endpoints. Integrating this component into their endpoint management system enables administrators to easily migrate client systems and manage rollout activities using Dell Client Manager, initiate scans and other troubleshooting- and health-related tasks from the Dell Client Manager console, and view Symantec Endpoint Protection–based environments from a single console to help simplify reporting and remediation tasks.

Comprehensive, cost-effective, easy-to-manage endpoint protection

Symantec Endpoint Protection enables organizations to implement an integrated solution that protects from threats on multiple levels. And even though it provides a comprehensive array of endpoint protection features, Symantec Endpoint Protection still gives administrators the flexibility to scale their protection over time. They can start with a limited set of features and then enable additional capabilities as needed. Symantec Endpoint Protection can even be configured to work alongside other vendors' technologies, so that organizations can implement and configure the solutions they need to address specific requirements.

In addition, the integrated Symantec Endpoint Protection agent is designed for low memory and resource usage on notebooks, desktops, and servers, helping eliminate the administrative overhead and costs associated with multiple security products. Administrators can tune the agent to help maintain endpoint performance, reducing its resource usage during periods of high user activity.

To combat sophisticated, stealthy, and targeted attacks, organizations can no longer rely solely on traditional antivirus and anti-spyware software. Effective endpoint security requires implementing a holistic solution that can proactively protect against threats at multiple levels while providing seamless interoperability, simplifying management, and reducing total cost of ownership. Through its integration of advanced security technologies into a single multilayered agent, Symantec Endpoint Protection can meet all of these requirements—helping simplify security administration, save time and money, and provide high levels of endpoint security to protect critical enterprise assets.

Lauren Duda *is a product marketing manager for the Endpoint Security team at Symantec. Lauren currently has global product marketing responsibility for Symantec AntiVirus™ and Symantec Client Security software. She has a B.A. from the University of California, Los Angeles, and an M.B.A. from California State University, Long Beach.*

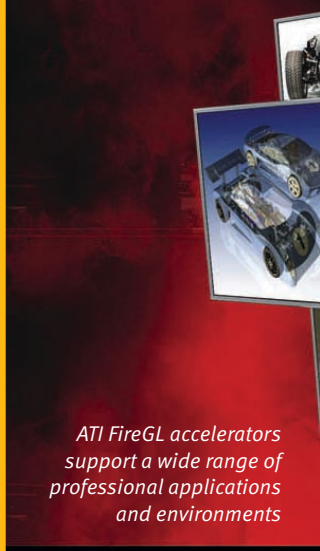
MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINK

Symantec Endpoint Protection:
www.symantec.com/endpoint

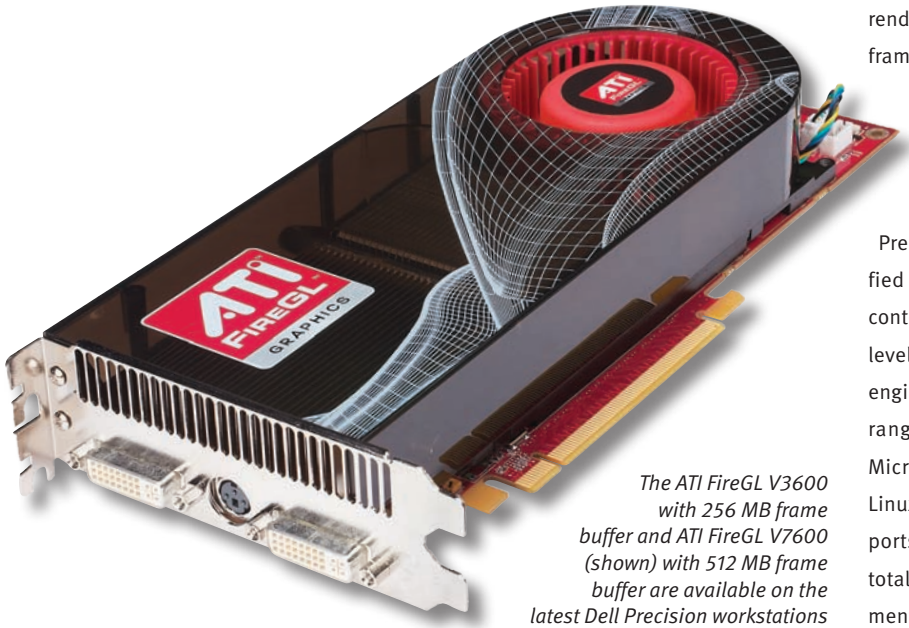
MAXIMIZING PRODUCTIVITY WITH ATI FIREGL WORKSTATION GRAPHICS ACCELERATORS

Offering innovative technology, a high level of reliability, and vibrant visual fidelity, ATI FireGL™ graphics accelerators from AMD are designed to enhance 3D application performance and increase user productivity in professional environments.



ATI FireGL accelerators support a wide range of professional applications and environments

The latest series of ATI FireGL workstation graphics accelerators from AMD are designed to provide comprehensive high-performance graphics capabilities for the most demanding professional-level 3D applications. Incorporating options ranging from a 256 MB frame buffer in entry-level cards to a 2 GB frame buffer in ultra-high-end models, these accelerators help boost both performance and user productivity.

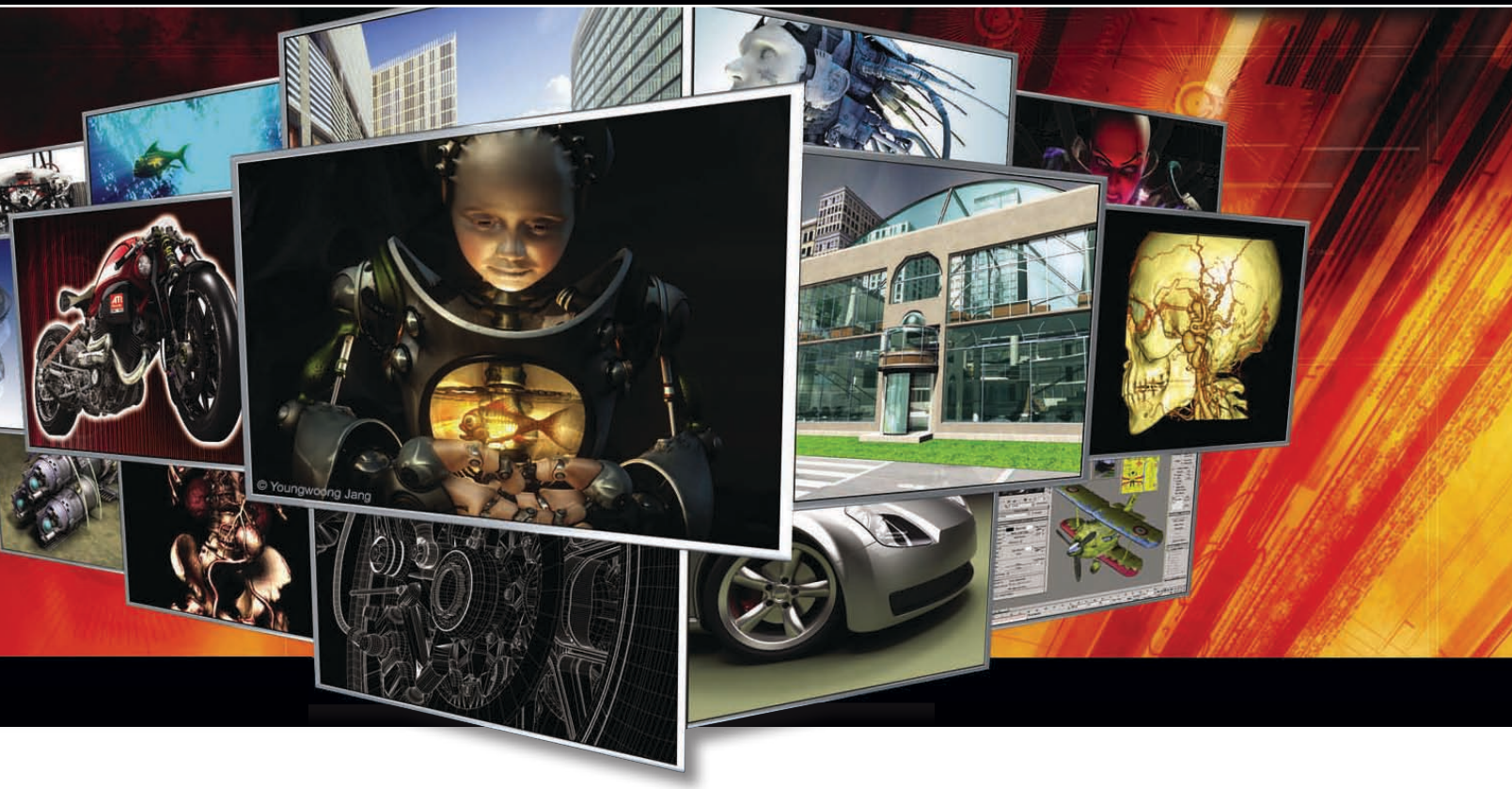


The ATI FireGL V3600 with 256 MB frame buffer and ATI FireGL V7600 (shown) with 512 MB frame buffer are available on the latest Dell Precision workstations

Based on a next-generation unified shader architecture, and integrating up to 320 processors in the graphics processing unit (GPU), these workstation cards help maximize throughput by automatically directing graphics power where it is needed when running applications built on Microsoft® DirectX® 10 or OpenGL 2.1 technology. This new series of products intelligently manages computational resources to enhance GPU utilization and enable real-time rendering of complex models and scenes, while increasing frame rates when animating.

Innovation and reliability from a technology leader

ATI FireGL accelerators, available as an option in Dell Precision™ workstations, are thoroughly tested and certified with major computer-aided design (CAD) and digital content creation (DCC) applications, helping ensure a high level of reliability. In addition, the ATI FireGL line has been engineered to deliver innovation and reliability for a wide range of professional environments, including the Microsoft Windows® XP, Microsoft Windows Vista™, and Linux® operating systems. The unified driver, which supports all ATI FireGL workstation products, helps reduce total cost of ownership by simplifying installation, deployment, and maintenance.



ATI FireGL accelerators also incorporate AMD's innovative AutoDetect technology. As users open new 3D applications, or move between them, optimized ATI FireGL graphics driver settings can automatically update the configuration to help maximize performance regardless of the user's workflow demands.

To further leverage these accelerators, users can take advantage of the massive stream computing capabilities of the GPU to accelerate compute-intensive tasks for software such as physics, structural analysis, and fluid dynamics applications.

Vibrant visual fidelity

Designed with a 10-bit display pipeline and support for high dynamic range output, ATI FireGL accelerators can produce over 1 billion colors for vibrant visual fidelity. These next-generation cards feature two dual-link-enabled Digital Visual Interface (DVI) outputs that can drive 30-inch Dell UltraSharp™ LCD panels and are capable of generating a multi-monitor desktop more than 5,000 pixels wide.

Native multi-card support also enables users to drive up to four displays by utilizing two ATI FireGL cards in a single workstation. Both high-definition component output and quad-buffered stereoscopic 3D output are available on high-end and ultra-high-end models, offering additional levels of realism for specialty applications.

For more information on ATI FireGL workstation graphics accelerators for Dell Precision workstations, visit ati.amd.com/firegl/dell.

Now with multi-card support, two ATI FireGL accelerators can drive four displays in a single workstation

ADDRESSING WEB-BASED APPLICATION PERFORMANCE AND AVAILABILITY WITH CITRIX NETSCALER

BY JOHN GUDMUNDSON

The ubiquity of the Web simplifies many aspects of delivering application services. However, the inherent performance and security inefficiencies of networking protocols can negatively affect user experience. Citrix® NetScaler® application delivery controllers provide advanced application-level acceleration, availability, and security functionality to address these issues while helping reduce the cost of delivering these applications.

Related Categories:

Application development

Application servers

Citrix Systems

Performance

Transmission Control Protocol (TCP)

Web applications

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

Since the rise of the Internet, enterprises have done an excellent job of providing network-level infrastructure and security. However, an increasingly mobile workforce, globalization, and a growing dependency on Web-based applications are illuminating problems that a network-only focus cannot address. Application users need quick response times, high availability, and application-layer security for the mission-critical Web-based applications used today. Citrix NetScaler application delivery controllers can address these challenges and provide advantages to a broad range of applications from leading independent software vendors.

Identifying Web-based application challenges

Routers, switches, firewalls, and other networking infrastructure, along with advanced dynamic routing protocols, have created an efficient method of moving packets. Optimal paths are selected, networking protocols can retransmit dropped packets, and "connection-oriented" protocols help ensure packets are eventually delivered. Yet, for all the networking advances that have been made in Open System Interconnection (OSI) Layers 2, 3, and 4, some issues persist when this infrastructure is used for high-performance applications communicating across wide area networks (WANs). Moreover, many of today's most pressing availability, performance, and security issues require Layer 7 visibility and processing.

TCP/IP, the protocol that Web-based applications are built on, originated with the premise of ensuring connectivity under any circumstances. HTTP was initially designed for academics to share research papers. Performance and security, especially for highly interactive, real-time, and security-sensitive applications, were not initial design points for the underlying infrastructure Web-based applications rely on.

Even using the fastest routers with wire-speed throughput and optical transport links, communicating to a global user base can add hundreds of milliseconds of latency to packet transport. The TCP/IP emphasis on availability magnifies the impact of this latency, in many cases increasing network congestion and frequently degrading application performance. Often, this delay causes network- or application-layer time-outs and forces traffic regeneration, which only further exacerbates the problem.

Only so much data (often less than 64 KB) can be transmitted before the receiver must send an acknowledgment to the transmitter. Although very few packets are typically dropped, these dropped packets force a retransmission of some packets previously sent successfully. Data rates from transmitters must frequently be throttled back because of wire congestion or the inability of the receiver (or routers along the way) to buffer the packets. When this throttling occurs, transmission rates only slowly return to normal

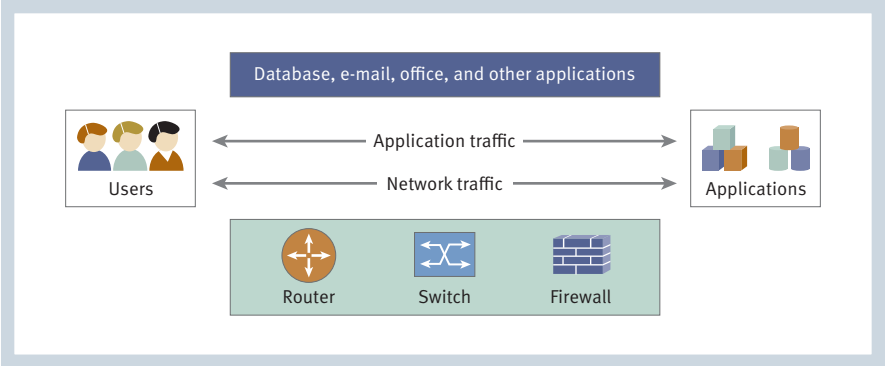


Figure 1. *Application and network traffic*

levels. The net result is that a connection specified as 100 Mbps, for example, cannot carry more than a small fraction of this load. In many instances, a 45 Mbps T3 link with 90 milliseconds of round-trip time provides barely over T1-level throughput.

HTTP, the application protocol the Web is built on, is notoriously chatty. This inherent chattiness, when combined with the highly interactive nature of Web and Web 2.0 applications, exacerbates the inherent performance limitations of the TCP architecture, and has created inefficiencies when scaled globally. Applications originally meant to be used within an office, but now used globally, suffer from poor performance. Because the core problems are associated with essential infrastructure, rewriting applications, increasing bandwidth, and adding servers fail to resolve them.

Addressing networking and application challenges

Solving these problems requires addressing the core networking behaviors that do not align with the behavior of Web-based applications. Addressing these behaviors frequently requires operating not only at the networking level, but also at the application layer (see Figure 1). Because the packet delivery infrastructure is opaque to the applications themselves, a viable solution must have insight into, and be transparent to, both elements.

Citrix NetScaler meets this requirement by integrating traditional server and network load-balancing services with the application-aware functionality necessary to efficiently deliver highly interactive Web-based applications across the globe. Specifically, NetScaler incorporates extensive network and application-layer capabilities such as protocol optimizations, compression, caching, and application security to accelerate application performance, increase application availability, and enhance application security (see Figure 2).

Application acceleration

Citrix NetScaler incorporates networking and application techniques to accelerate application performance. To offset the negative consequences of the original TCP/IP architecture, NetScaler provides numerous networking optimizations that help minimize the impact of network latency, reduce network congestion, and ultimately increase application performance. Examples include the following:

- **TCP windows scaling:** Enables administrators to set the TCP window size at greater than 65 KB
- **Selective acknowledgment and retransmission:** Intelligently manages data acknowledgments and retransmits dropped packets, helping reduce wire congestion
- **TCP fast ramp:** Overcomes the standard TCP slow-start algorithm by initially setting the transmission rate equal to what the client can receive, helping increase the efficiency of bandwidth use and reduce response times
- **Client keep-alive:** Intelligently determines when a client connection can be kept open and when it must be closed; maintaining client connections helps reduce network congestion and response times

Using compression also helps accelerate data delivery. Compressing data helps reduce traffic, thereby minimizing the likelihood of packet loss and the number of data acknowledgments while increasing data transfer efficiency and application performance. By offloading this processing-intensive data compression task, NetScaler appliances enable servers to operate with increased efficiency. NetScaler is designed to compress data by three to five times, depending on content type, thus helping reduce bandwidth needs.

Content caching is designed to further reduce user response times. Multiple users often request the same information, which may

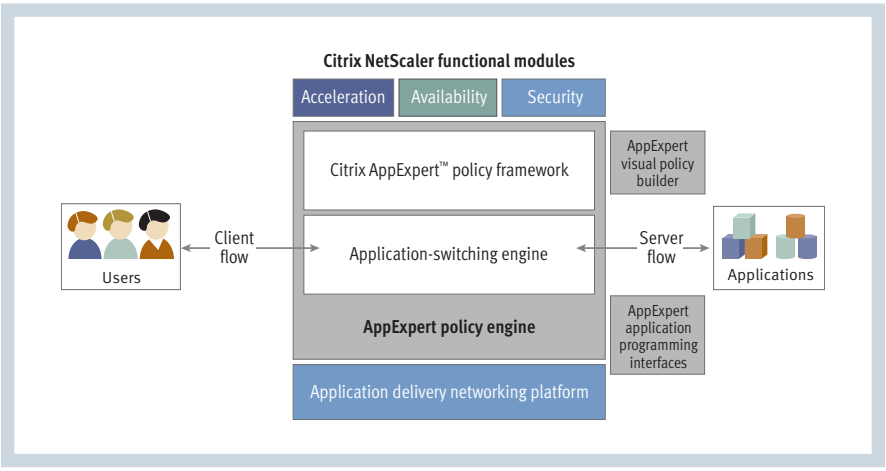


Figure 2. *Citrix NetScaler application delivery controller architecture*

be static or dynamically generated by the application. By caching data ahead of time on the appliance, NetScaler enables an immediate response, further enhancing user experience.

Application availability

Functionality such as load-balancing, content-switching, and offload techniques helps ensure servers are reachable. When servers are load balanced, each server gets its proper share of requests, helping maximize its efficiency and ability to respond in a timely manner. Load-balancing methods can be as simple as round-robin algorithms or as complex as load distribution based on real-time processor utilization rates. Health checks help ensure servers and databases are operational before requests are forwarded to them. If a session demands a persistent connection, it can be maintained by a variety of methods, including Secure Sockets Layer (SSL) IDs, cookies, or tokens created from an aspect of the initial request's payload—for example, an application-level transaction ID within a Simple Object Access Protocol (SOAP) request.

With content switching based on application-layer content, load balancing is further enhanced as traffic is directed to servers optimized for specific content requests. For example, requests for dynamic content (such as URLs with a suffix of .asp, .dll, or .exe) can be directed to one server or farm, and requests for static content directed to another. Organizations can achieve excellent server farm efficiencies without duplicating content across servers.

The multiple offloading methods Citrix NetScaler provides help increase server availability and reduce application response times. SSL encryption, especially given the increased use of large key sizes and complex encryption algorithms, is incredibly processor intensive for application and Web servers. Offloading SSL handling onto purpose-built appliances frees servers to serve content, generally leading not only to reduced server processor utilization, but also to increased application availability and accelerated response times.

Workflow	Test	Response time (seconds)	Relative decrease in response time with NetScaler
MySite	Baseline	6.65	47.7%
	NetScaler	3.48	
Search Query	Baseline	5.24	26.7%
	NetScaler	3.84	
Document Library	Baseline	6.49	48.2%
	NetScaler	3.36	
Team Site	Baseline	6.29	82.0%
	NetScaler	1.13	

Note: Response time based on configuration with 512 Kbps bandwidth and 140 ms latency.

Figure 3. *Test results when running Microsoft Office SharePoint with Citrix NetScaler*

The basic compute-intensive TCP/IP session setup and teardown process is also largely offloaded from the server. NetScaler terminates client TCP connections and, in turn, manages a set of long-lived sessions with the servers. This multiplexing is designed to reduce the number of TCP sessions a server must maintain by 100 times or more.

Application security

Networking-level security holes have largely been plugged. However, the applications themselves remain exposed, because the network must allow users to reach these applications. Application-level attacks take advantage of seemingly innocuous coding flaws.

Application-layer protection helps mitigate the application-layer attacks prevalent today. Network security cannot prevent attacks based on SQL injection, buffer overflows, form field tampering, or cross-site scripting. Moreover, these threats are also extremely difficult to mitigate using intrusion prevention systems.

To help deliver zero-day protection against these exploits and many more, Citrix NetScaler employs the positive security model. Rather than relying on signatures, this model combines knowledge of HTTP standards and protection against known application attack vectors with the ability to automatically learn an application's correct behavior. Abnormal application behavior is treated as potentially dangerous and, depending on administrative preference, blocked and/or logged. The appliance automatically learns legitimate behavior and is designed to block known malicious usage without false positives.

Testing performance enhancements with Citrix NetScaler

Citrix NetScaler has demonstrated enhancements that not only address traditional server availability concerns, but also help increase application performance. Testing done in conjunction with Microsoft, among many other software vendors, has shown tangible advantages. Such independent testing occurred on-site in the vendor's laboratories. For example, when running Microsoft® Office SharePoint® software on Dell™ PowerEdge™ servers in conjunction with NetScaler, Microsoft showed up to an 82 percent reduction in latency for various workflows (see Figure 3).[1]

Helping applications run as intended

Applications once all but useless because of the inherent limitations of legacy protocols and WAN distances can excel when deployed with Citrix NetScaler. Using Citrix NetScaler can help organizations increase performance levels and server availability, maximize operational efficiency, and enhance security all the way to the applications themselves.

John Gudmundson *is a senior product marketing manager in the Applications Networking Group at Citrix Systems.*

[1] For the complete report, see "Microsoft SharePoint and Citrix Application Optimization Deployment Best Practices and Performance Validation," by Citrix Systems, May 2007, www.hqfastapps.com/HQFastApps/MicrosoftSharepointandNetScalerWANScalerDeploymentandPerformanceValidationv10.pdf.

FLEXIBLE MANAGEMENT TOOLS FOR DELL REMOTE CONSOLE SWITCHES

BY MAX A. BENHAM
WILLIAM MUSCATO

The latest firmware updates for Dell™ 2161DS-2 and 4161DS remote console switches enhance remote server management through an on-board Web browser–based interface, virtual media features, and support for Avocent® KVM (keyboard, video, mouse) over IP technology and Avocent DSView® 3 management software.

Related Categories:

Keyboard, video, mouse (KVM)

Remote management

Storage management

Virtual media

Visit ***DELL.COM/PowerSolutions*** *for the complete category index.*

Availability and security are cornerstones of data center operations and maintenance. Network administrators continue to require access to and real-time control over their servers and other IT assets while maintaining a largely automated environment. The latest firmware updates for Dell 2161DS-2 and 4161DS remote console switches help meet these remote management needs through Avocent KVM (keyboard, video, mouse) over IP technology in conjunction with Avocent DSView 3 management software.

Understanding KVM over IP technology

KVM over IP digitizes keyboard, video, and mouse data and can then send it over TCP/IP connections. It can connect directly to a server using existing network infrastructure, does not require adding software or hardware to computers, and supports both local and remote users. It is designed to work in heterogeneous hardware environments and is well suited for managing multilocation data centers and branch offices.

Utilizing Dell 2161DS-2 or 4161 remote console switches with Avocent KVM over IP technology requires two components: the switch and a server interface pod (SIP). The hardware consists of a rack-mountable KVM switch that integrates the traditional functionality of analog KVM switches with the digital technology of KVM over IP. The switches can use 1U of rack space or be mounted in the 0U area of a Dell rack. Each switch can manage up to 128 servers, or connect to Dell analog KVM switches to manage up to 256 servers. By combining analog and digital technology, these switches can provide flexible, centralized server control while helping significantly reduce cable volume and enabling secure remote access.

The local port provides the analog connection for server access. This connection accommodates access for either a data center crash cart or a rack-mounted KVM solution. A 10/100/1,000 Mbps Ethernet interface provides the digital remote connection for server access, digitizing the server image and then transporting it through a TCP/IP connection across a corporate intranet or over the Internet to a remote user.

For more information about the advantages of KVM technology, see the "Key advantages of KVM technology" sidebar in this article.

Integrating Dell KVM console switches with Avocent DSView3 software

The latest firmware updates for Dell 2161DS-2 and 4161DS remote console switches are designed to broaden and deepen the reach of KVM over IP infrastructure control for network administrators. The enhancements include the following:

- An on-board Web browser–based interface that allows direct access to the switch
- A virtual media implementation within the switch that enables out-of-band file transfers as well as application and OS patch deployment
- Integration with Avocent DSView 3 software, which allows administrators to control switch power and manage other devices from a single console

Figure 1 shows the basic elements of a data center environment incorporating Dell remote console switches, Cyclades® ACS advanced console servers from Avocent, Cyclades PM intelligent power distribution units (IPDUs) from Avocent, and an Avocent MergePoint™ 5200 service processor (SP) manager. In addition, this figure shows the three tools through which administrators can access such an environment: Dell Remote Console Software (RCS), the remote console switch Web browser–based interface, and Avocent DSView 3.

Using Dell remote console switches in conjunction with Avocent KVM over IP technology and Avocent DSView 3 software can offer multiple advantages, including simplified control over multiple systems, easy integration into existing infrastructures, comprehensive management capabilities, increased efficiency and security, and enhanced availability.

Simplified control over multiple systems

Dell remote console switches support several methods of remotely managing, viewing, and controlling attached servers:

- The on-board Web browser–based interface, which provides a secure connection to individual switches using a standard Web browser
- Dell RCS, which runs on a Java platform and enables administrators to manage multiple switches at once
- Avocent DSView 3 management software, which offers comprehensive data center control and enhanced security

Each of these options provides access to the system during reboot, even at the BIOS level, without requiring software installation on the target browser. To transmit KVM signals, Dell RCS uses industry-standard TCP/IP connections and confirms 128-bit Secure Sockets Layer (SSL) encryption. Administrators can also select Data Encryption Standard (DES) or Triple DES (3DES) encryption.

Easy integration into existing infrastructures

SIPs connect to the KVM port of target servers and send analog signals over a Category 5 (Cat 5)

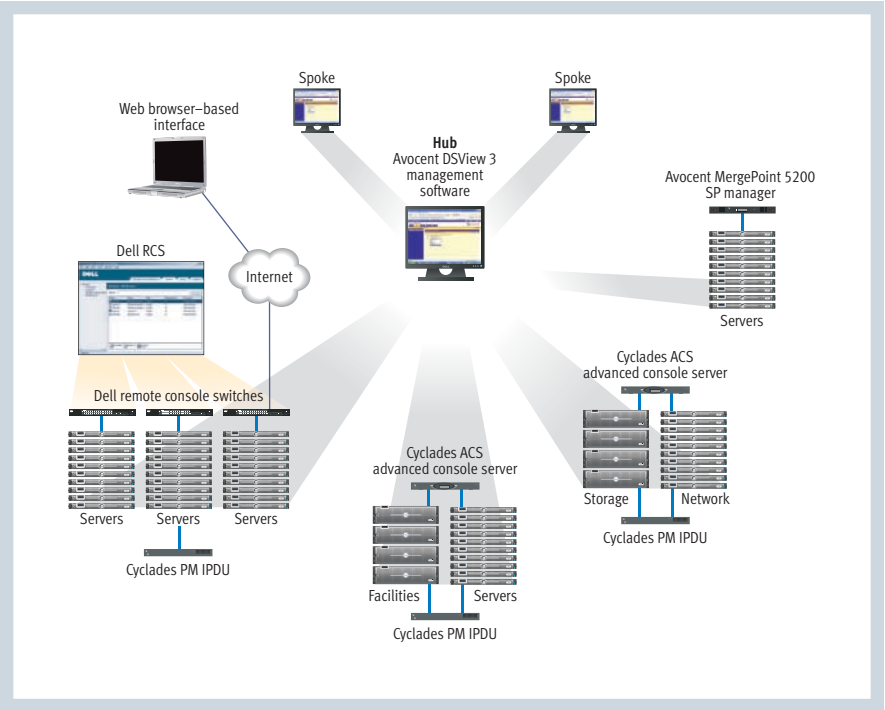
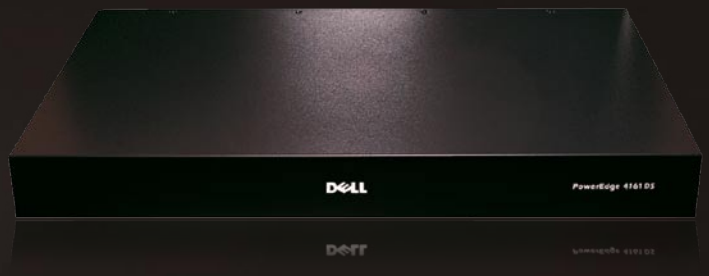


Figure 1. *Example environment incorporating Dell remote console switches, Avocent KVM over IP technology, and Avocent DSView 3 software*

KEY ADVANTAGES OF KVM TECHNOLOGY

KVM infrastructures can provide multiple advantages in enterprise environments, including the following:

- **Centralized management:** Switches utilizing KVM over IP technology provide access to servers and devices from a central location that administrators can easily access remotely.
- **Out-of-band server access:** Because they provide a direct connection to keyboard, video, and mouse ports, KVM switches enable BIOS-level access independent of servers or networks—meaning administrators can still manage servers even if the network fails.
- **Increased efficiency and reduced costs:** By providing centralized access to IT infrastructures, KVM technology enables administrators to centralize expertise, increase response times, and reduce the cost and time of physically going to the server room.
- **Increased data center security:** KVM technology enables administrators to maintain servers in a secure environment while still enabling full access to those servers. It also provides additional security such as password control and detailed logging of user access to servers.
- **Increased responsiveness and reduced downtime:** With KVM technology, administrators can respond immediately to problems from any location. Expert centralized staff can immediately diagnose and repair problems in remote data centers and branch offices.
- **Reduced power and space requirements:** By rendering multiple sets of redundant peripherals unnecessary, KVM technology helps enterprises reduce power and space requirements.



MANAGE UP TO 256 SERVERS ANYTIME, ANYWHERE

With Dell Remote Console Switches you can remotely manage, view, and control your servers. Each switch can manage up to 128 servers, or connect to Dell analog KVM switches to manage up to 256 servers. Get real-time control—anytime and anywhere.

SIMPLIFY SERVER MANAGEMENT AT DELL.COM

"By combining analog and digital technology, Dell 2161DS-2 and 4161DS switches can provide flexible, centralized server control while helping significantly reduce cable volume and enabling secure remote access."

unshielded twisted-pair (UTP) cable back to the switch. Designed to reduce cable volume, SIPs can help save rack space and allow administrators to position servers more than 50 feet from the switch.

Comprehensive management capabilities

In the past, switch administrators have had access to and control over servers. When servers failed or when hub, router, or headless server problems arose, however, they could not rely on their KVM switch infrastructure to resolve them.

The latest firmware updates for Dell remote console switches add support for Avocent DSView 3 software, providing enhanced management capabilities and allowing administrators to manage additional devices from their KVM management console. Specifically, administrators can now implement Avocent MergePoint SP managers to utilize the Intelligent Platform Management Interface (IPMI) software in eighth- and ninth-generation Dell PowerEdge™ servers. They can also implement serial console switches to manage hubs, routers, and headless servers as well as IPDUs. Incorporating this hardware into their Dell KVM over IP infrastructure can provide administrators with comprehensive out-of-band access and control over heterogeneous data center environments.

Increased efficiency and security

Physically getting up and going to the server room to install a patch or run diagnostics costs administrators valuable time, and the increased foot traffic through the data center can also increase exposure to potential security risks. The virtual media capability added in the latest firmware updates to Dell remote console switches is designed to address these problems. By enabling administrators to map CDs and other storage media to a remote server to perform file transfers, application and OS patches, and diagnostics, virtual media helps increase efficiency without compromising physical security.

Enhanced availability

Failed servers can cause major difficulties when administrators are working remotely or in environments with tightly controlled physical access. Dell remote console switches and Avocent KVM technology provide administrators with multiple ways to resolve server problems in these types of environments.

If the failed server is an eighth- or ninth-generation Dell PowerEdge server with IPMI 2.0 enabled, administrators can remotely view the system event log to help identify the problem. They can then use the Avocent MergePoint SP manager to power down or power cycle the server if necessary. If the server does not support IPMI, administrators can use Avocent DSView 3 to access the IPDU infrastructure and power cycle the server.

If a problem occurs with a server that uses a serial console as its management port, administrators can access the serial console switch through DSView 3 to help identify the problem. They can then use either the command-line interface to reboot the device or the IPDU infrastructure to power cycle the server, and view the server reboot using the KVM over IP connection.

Simplifying management in KVM infrastructures

Dell 2161DS-2 and 4161DS remote console switches coupled with Avocent KVM over IP technology and Avocent DSView 3 software can provide a single console for cost-effective out-of-band access, management, and power control—helping reduce mean time to repair while maintaining the security of the physical environment. Implementing these elements in data centers utilizing eighth- and ninth-generation Dell PowerEdge servers enables administrators to enhance their out-of-band management capabilities, increase efficiency, and simplify data center management.

Max A. Benham *is the Dell appliance account manager at Avocent. Previously, he led the Avocent Original Equipment Manufacturer (OEM) Program Management organization. Max has a B.S. in Economics and a B.A. in Slavic Languages and Literature from the University of Washington.*

William Muscato *is a Dell product marketing manager for rack and rack infrastructure products. He has previously held executive positions in the technology, finance, and travel industries, providing business, sales, and product management expertise in emerging technologies and online services. William has an engineering degree from the Rochester Institute of Technology and an M.B.A. from the Ohio State University.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINK

Avocent:
www.avocent.com

CONFIGURING INFINIBAND ON HPC CLUSTERS USING THE OPENFABRICS ENTERPRISE DISTRIBUTION

BY MUNIRA HUSSAIN
ARUN RAJAN
SREERAM VEDANTHAM
JACOB LIBERMAN

The OpenFabrics Enterprise Distribution (OFED) provides a standardized software stack for high-performance computing (HPC) clusters using the InfiniBand interconnect. This article describes the development and architecture of OFED along with test results measuring the bandwidth and latency of OFED-based HPC clusters of Dell™ PowerEdge™ servers.

Related Categories:

Dell PowerEdge servers

High-performance computing (HPC)

InfiniBand

OpenFabrics

Visit ***DELL.COM/PowerSolutions*** *for the complete category index.*

The InfiniBand interconnect is designed to deliver high-bandwidth, low-latency Interprocess Communication (IPC) between high-performance computing (HPC) cluster nodes. It initially supported single data rate (SDR) network interface cards (NICs) on PCI Extended (PCI-X) buses, and later added support for SDR on PCI Express (PCIe) buses. With the advent of double data rate (DDR) InfiniBand NICs and switches, InfiniBand has advanced to offer increased unidirectional bandwidth of 20 Gbps or higher. InfiniBand uses the 8B/10B encoding mechanism to enhance signal transmissions.

InfiniBand is based on a switched fabric architecture of serial point-to-point links, where each point could be a direct interface to a host system or channel adapter. These links can also connect to either host channel adapters (HCAs), used primarily in servers, or target channel adapters (TCAs), used primarily in storage subsystems. Features such as zero-copy and Remote Direct Memory Access (RDMA) help reduce processor overhead by directly transferring data from sender memory to receiver memory without involving host processors.

As InfiniBand has grown in popularity among open source communities and scientific labs, multiple versions of the InfiniBand stack have become available. The need for a standardized stack led to the creation of the OpenIB Alliance, now known as the OpenFabrics Alliance, which has created the OpenFabrics Enterprise Distribution (OFED). This article outlines how OFED has evolved as well as key components of its architecture, and describes test results measuring the bandwidth and latency of OFED-based HPC clusters of Dell PowerEdge servers.

Development and architecture of the OpenFabrics Enterprise Distribution

Because InfiniBand hardware supports a wide array of features, InfiniBand software—including both kernel- and user-level components—can be extremely complicated. Fully functional InfiniBand software stacks comprise many libraries, application programming interfaces (APIs), protocols, and modules. Rather than attempting to manage this complexity by developing competing InfiniBand software implementations, hardware vendors, research institutions, and governmental organizations worked together to form the OpenIB Alliance, now known as the OpenFabrics Alliance—a nonprofit group committed to the development and proliferation of open source, transport-independent software for high-performance server and storage connectivity. The group's mission statement includes three directives: to "unify the cohesive development of a single open-source, RDMA-enabled, transport independent software stack that is architected for high-performance, low-latency, and maximized efficiency"; to "promote industry awareness, acceptance, and benefits of these solutions for server and storage clustering and connectivity applications"; and to "manage the interoperability testing and certification of the software running on

different hardware solutions."[1] Members of the OpenFabrics Alliance include hardware vendors such as Dell, Intel, AMD, and Cisco, and governmental organizations such as Sandia National Laboratories and Los Alamos National Laboratory.

Since 2004, the OpenFabrics Alliance has achieved several milestones. Although the organization originally focused on developing a standardized software stack for InfiniBand on the Linux® OS, in 2005 the group also began supporting InfiniBand on the Microsoft® Windows® OS. In 2006, the group added support for the Internet Wide Area RDMA Protocol (iWARP), an RDMA-enabled transport based on the Ethernet specification. Finally, the OpenFabrics Alliance has also enjoyed success in its primary goal of creating a standardized software stack for InfiniBand on Linux, with the InfiniBand kernel module now adopted in the upstream Linux kernel.

Evolution of the OpenFabrics Enterprise Distribution

In 2006, the OpenFabrics Alliance released the first revision of OFED, a commercial-quality software distribution that supports clustering and grid connectivity using InfiniBand. OFED bundles the latest InfiniBand kernel modules available in the OpenFabrics source code repository along with additional applications and modules such as Message Passing Interface (MPI) software, performance benchmarks, and communication protocol drivers. Because OFED is the result of open source community development, it is not tied to a particular hardware or software vendor, enabling cluster users to deploy InfiniBand hardware from various vendors with a single generic software stack.

OFED 1.0 contained the OpenFabrics core HCA drivers as well as upper-level protocols such as IP over InfiniBand (IPoIB), Sockets Direct Protocol (SDP), a SCSI RDMA Protocol (SRP) initiator, and User Direct Access Programming Library (uDAPL). It also included additional OpenFabrics utilities such as the OpenSM Subnet Manager along with various diagnostic tools and performance benchmarking utilities, and was bundled with the Ohio State University (OSU) MPI implementation and the Open MPI stack supporting the InfiniBand interface.

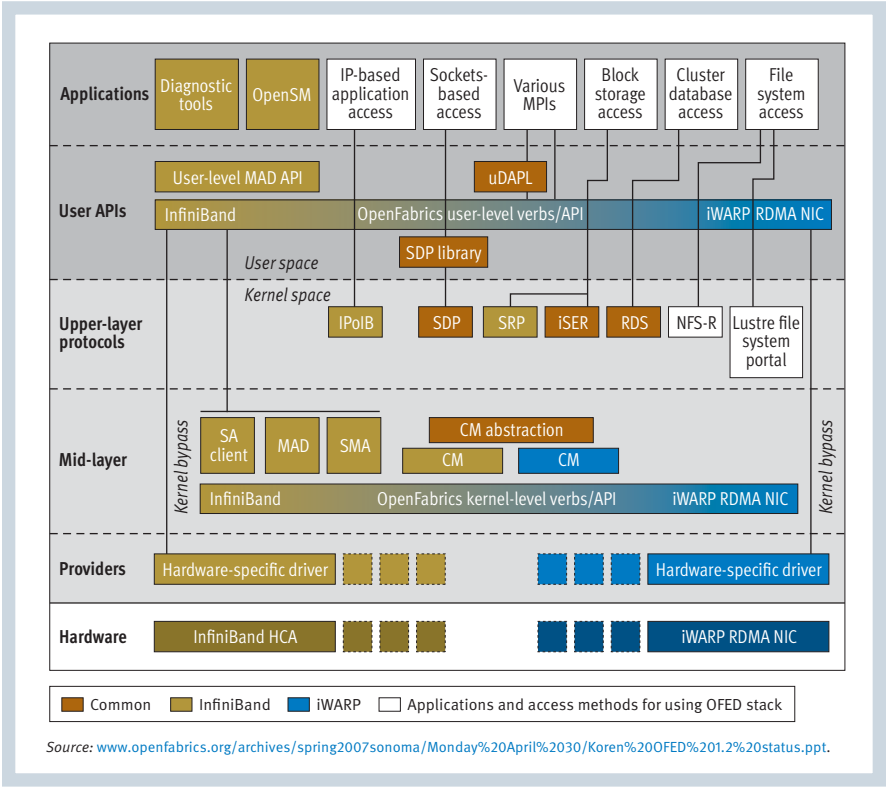
OFED 1.1 added hardware support and support for updated MPI packages. This version also provided drivers and firmware for 20 Gbps DDR InfiniBand HCAs and memory-free HCAs, included an updated version of Open MPI, and replaced the OSU MPI implementation with MVAPICH.

OFED 1.2, the most recent version, contains several significant enhancements over previous versions. First, it includes an enhanced installation and removal tool that helps significantly simplify the deployment of OFED drivers and utilities over existing installations. A single ofedinstall wrapper script can remove or back up previous versions of InfiniBand drivers and modules, helping avoid redundancies. OFED 1.2 also incorporates additional IPoIB features, including support for IPoIB connected mode (as opposed to datagram mode) and an IPoIB high-availability driver based on the Linux Ethernet bonding driver. Furthermore, OFED 1.2 includes updated versions of Open MPI and MVAPICH as well as added support for PGI compilers. Open MPI 1.2 also supports shared memory on dual- and quad-core processor sockets.

The evolution of OFED reflects the OpenFabrics Alliance commitment to developing vendor-neutral commercial-grade software for RDMA-enabled transport-based interconnects while helping support the adoption of best-in-breed HPC management utilities and applications.

OpenFabrics Enterprise Distribution software stack and components

OFED is the result of efforts by the OpenFabrics Alliance Enterprise Working Group to create a standard software stack distribution with the participation of multiple vendors. OFED 1.2 includes kernel- and user-level components comprising multiple tiers, as shown in Figure 1.



Source: www.openfabrics.org/archives/spring2007sonoma/Monday%20April%2030/Koren%20OFED%201.2%20status.ppt.

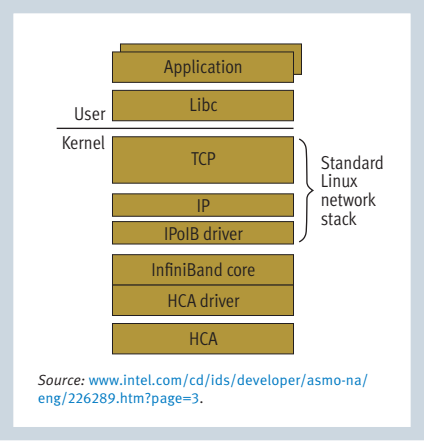
Figure 1. *OpenFabrics Enterprise Distribution 1.2 architecture*

[1] "OpenFabrics Overview," by the OpenFabrics Alliance, January 2007, www.openfabrics.org/downloads/about%20us-long.pdf.

The kernel space is a three-tier layer, with the hardware-specific drivers (which could be InfiniBand HCA drivers or RDMA NIC drivers) forming the bottommost tier. The mid-layer, also known as the core layer, consists of InfiniBand verbs, which are provided by hardware drivers translating the semantic functions into kernel-level APIs for the OS. This tier could be thought of as middleware between the kernel API and the associated HCA driver. It also comprises the General Services Managers—a group created by the Subnet Manager—and their respective agents, which are responsible for the management and operation of InfiniBand hardware and the routing of data. General Services Managers and agents include the Subnet Administrator (SA) client, Management Datagram (MAD), the Subnet Manager Agent (SMA), the Communication Manager (CM), and the Performance Manager Agent (PMA; not shown).

In an OFED-based infrastructure, the Subnet Manager configures host devices by performing a topology discovery and interacting with host SMAs, which respond to subnet management packets. MAD, an interface for sending and receiving data, includes components such as the Subnet Manager Interface (SMI) and General Services Interface (GSI). The SMI provides a unique queue pair (QP) mechanism termed QP0 to send data to and receive data from the Subnet Manager. The GSI uses a unique QP1 mechanism that allows clients to send and receive data through various General Services Managers and their agents. The PMA communicates with a General Services Manager called the Performance Manager and responds to packets, in particular to check for hardware counter errors. The SA client is responsible for communication between clients and the SA.

The upper layer of the kernel space consists of protocol stacks such as IPoIB, SDP, SRP, Internet SCSI (iSCSI) Extensions for RDMA (iSER), Reliable Datagram Sockets (RDS), Network File System over RDMA (NFS-R), and the Lustre file system portal for InfiniBand. Drivers for these protocols also reside in this layer. The IPoIB stack allows the use of IP with InfiniBand (see Figure 2); the IPoIB driver is a standard Linux driver enabling applications to use this stack directly without modification. The SDP stack, used for socket stream semantics, utilizes RDMA with TCP/IP as well as a zero-copy feature to help increase performance. The SRP stack provides SCSI packet transfers over InfiniBand. iSER maps the iSCSI protocol over RDMA to enable zero-copy message transfers. Both SRP and iSCSI provide block storage access.



Source: www.intel.com/cd/ids/developer/asmo-na/eng/226289.htm?page=3.

Figure 2. *IP over InfiniBand architecture*

The user space is a two-tier layer that provides access from applications to user APIs, and uses InfiniBand verbs similarly to the kernel space. uDAPL provides the RDMA APIs to the user space. uDAPL, created by the Direct Access Transport (DAT) Collaborative, defines and standardizes a set of transport-independent, platform-independent APIs to help effectively utilize RDMA with InfiniBand and iWARP. An application built with uDAPL libraries can typically run on any RDMA-enabled fabric.

OFED 1.2 integrates Open MPI 1.2, an open source MPI implementation that combines the features of the MPI-1 specification and the MPI-2 extension to MPI-1. Open MPI 1.2 supports runtime interconnect selection through Modular Component Architecture (MCA) parameters, and includes the OpenIB MCA library, which links with the OFED stack at runtime to enable running jobs through InfiniBand.

OpenFabrics Enterprise Distribution performance on Dell PowerEdge servers

To test InfiniBand performance with OFED 1.2, in July 2007 Dell engineers performed benchmark tests on clusters of Dell PowerEdge servers using the Open MPI and MVAPICH libraries on both SDR and DDR InfiniBand. Open MPI is scalable and modular, and therefore can support multiple interconnects; its components also include runtime tuning parameters. MVAPICH supports the shared memory required for extended dual- and quad-core processors, and is efficient in its support for memory hierarchies in multi-core systems.

The test team used the Intel MPI Benchmarks (IMB) suite and the OSU MPI-level tests to evaluate unidirectional and bidirectional bandwidth and latency. These benchmarks are compiled with GCC, Intel, and PGI compilers with both the Open MPI and MVAPICH middleware libraries in the OFED 1.2 package.[2]

Cluster configurations in the test environment

The SDR InfiniBand–based cluster consisted of two Dell PowerEdge SC1435 servers with two dual-core AMD Opteron™ 2218 processors at 2.6 GHz and a 1 MB level 2 (L2) cache. Each node had 4 GB of memory using interleaved, fully buffered dual in-line memory modules (DIMMs) in each slot, as well as a single-port, memory-free Cisco SDR InfiniBand HCA (SFS-HCA-312-A1) connected to a PCIe riser in an x8 PCIe slot. Each node was connected to a 24-port Cisco SFS 7000P SDR InfiniBand switch. The cluster ran Platform Open Cluster Stack (OCS) 4.4.0 with the Cisco OFED 1.2 Roll and the Red Hat® Enterprise Linux 4 Update 4 OS with kernel version 2.6.9-42.EL.smp.

The DDR InfiniBand–based cluster consisted of two Dell PowerEdge 2970 servers with two dual-core AMD Opteron 2222 processors at 3.0 GHz and a 1 MB L2 cache. Each node had 8 GB of fully buffered DIMMS, as well as a single-port, memory-free Cisco DDR InfiniBand HCA (SFS-HCA-320-A1) connected to a PCIe riser

[2] For more information on these benchmarks, visit www.intel.com/cd/software/products/asmo-na/eng/307696.htm#mpibenchmarks and mvapich.cse.ohio-state.edu/benchmarks.

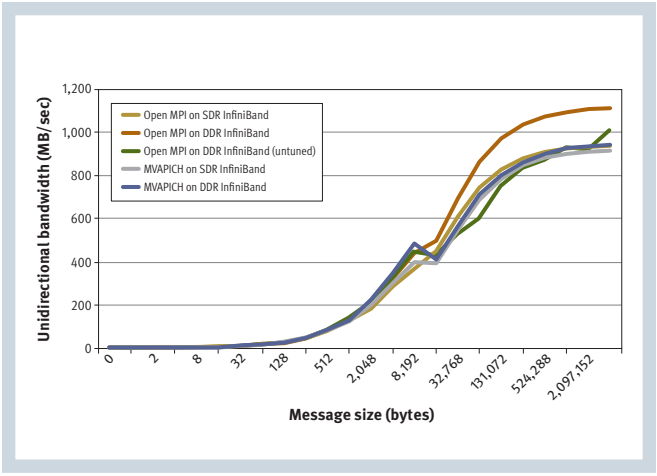


Figure 3. *Unidirectional bandwidth with OpenFabrics Enterprise Distribution 1.2, measured using the Intel MPI Benchmarks PingPong test*

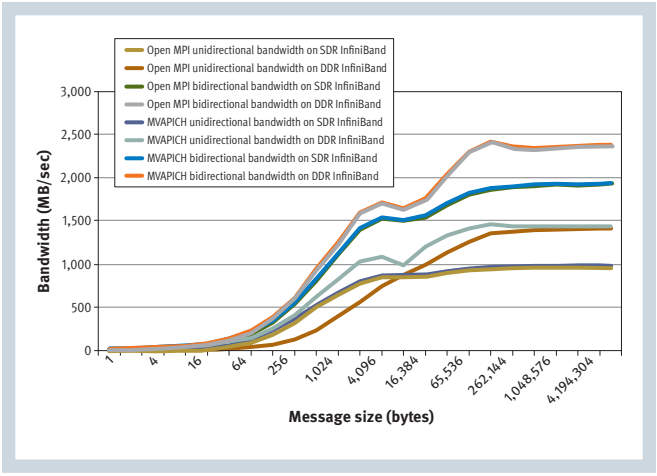


Figure 4. *Unidirectional and bidirectional bandwidth with OpenFabrics Enterprise Distribution 1.2, measured using the Ohio State University MPI bandwidth tests*

in an x8 PCIe slot. Each node was connected to a 24-port Cisco SFS 7000D DDR InfiniBand switch. This cluster ran the same software as the SDR InfiniBand–based cluster.

Test results: Bandwidth and latency

OFED 1.2 offers a variety of open source message passing libraries. Open MPI is scalable and modular, enabling it to support multiple interconnects, and its components incorporate runtime tuning parameters. MVAPICH, on the other hand, consists of both MPI1 and MPI2 implementations, and supports the shared memory necessary for extended dual- and quad-core processors. MVAPICH is efficient in providing memory hierarchy support in multi-core systems.

Figure 3 shows unidirectional bandwidth performance as measured using the IMB PingPong test. In this test, the sender node sends a message to the receiver node, and the receiver node then sends a message of the same size back to the sender node; the test team measured the bandwidth in unidirectional mode with 8B/10B encoding.

Figure 4 shows unidirectional and bidirectional bandwidth performance as measured using the OSU unidirectional and bidirectional bandwidth tests. When testing Open MPI with DDR InfiniBand, the test team used the command `-mca mpi_leave_pinned 1` to have the OS lock the memory and help ensure that messages remained in the same physical location until the HCA had finished transferring the data.

Figures 5 and 6 show latency as measured using the IMB PingPong test and OSU latency test, respectively. The test team measured latency as half of the total round-trip time. As Figures 5 and 6 show, all four configurations typically had extremely low latencies of between 3.5 and 4.2 nanoseconds.

OpenFabrics Enterprise Distribution integration with Platform Open Cluster Stack

Platform OCS, a comprehensive cluster computing software stack based on software developed by the San Diego Supercomputer Center (SDSC)

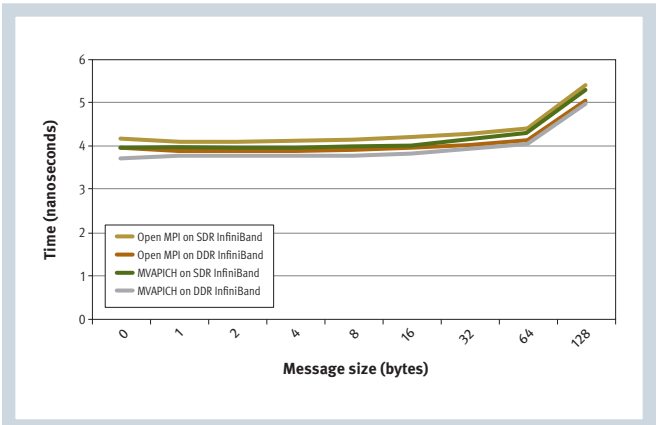


Figure 5. *Latency with OpenFabrics Enterprise Distribution 1.2, measured using the Intel MPI Benchmarks PingPong test*

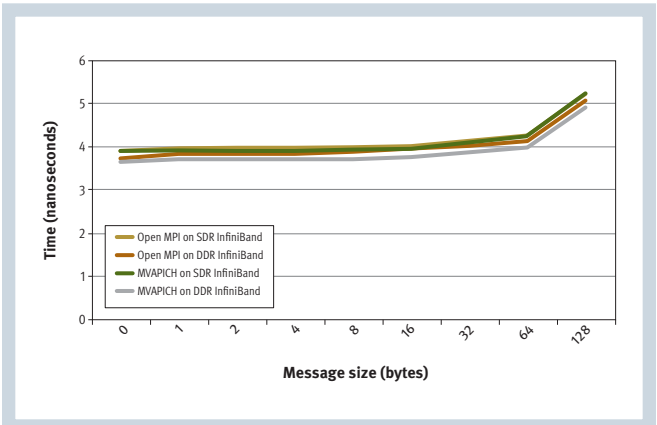


Figure 6. *Latency with OpenFabrics Enterprise Distribution 1.2, measured using the Ohio State University MPI latency test*

National Partnership for Advanced Computational Infrastructure (NPACI), is designed to simplify the deployment, maintenance, and management of large-scale Linux-based HPC clusters. Administrators can create OCS Rolls containing domain-specific software and applications, and third-party vendors can customize Rolls to meet specific requirements.

The latest version of the OCS stack, OCS 4.4.0, is based on Red Hat Enterprise Linux 4 Update 4. The Cisco OFED Roll (version 4.4.1), provided by Platform Computing as an add-on to the base OCS DVD, is a fully supported and verified addition to Dell HPC bundles that integrates the components from the Cisco OFED 1.2 stack necessary to install compute nodes with Cisco NICs and switches. It includes Cisco OFED drivers, utilities, libraries, and supported MPI stacks, and provides the necessary functionality to integrate OFED 1.2 with OCS. Administrators can use this Roll to perform tasks such as installing kernel modules, installing InfiniBand libraries, installing OFED NIC utilities, backing up conflicting Red Hat Package Manager (RPM™) files, and restoring RPM files to their original locations when uninstalling and upgrading InfiniBand HCA firmware. In addition to the standard OFED components, this Roll provides the hca_self_test utility, the Linpack binary compiled with Open MPI 1.2 (GCC version), and the latest firmware for Cisco HCAs.

Administrators can install the Cisco OFED Roll as part of a new front-end installation or on an existing front-end installation using the rollops tool. After completing the installation, they can follow the on-screen instructions should the system require them to log in again or to reboot. During installation, the Cisco-specific components, including the xhpl binary and some MPI examples, are installed in /opt/cisco_ofed. The standard OFED components are installed in /usr, with the MVAPICH and Open MPI 1.2 MPI implementations and MPI benchmarks in /usr/mpi and the InfiniBand diagnostic tools and utilities in /usr/bin. The Cisco OFED RPM packages are installed on all nodes, including the front end, regardless of whether an SDR or DDR card is present.

Previous OCS versions included Open MPI 1.1.4, but when the Cisco OFED Roll is installed, OCS uses Open MPI 1.2 instead. Because of Roll conflict, using both available versions of the InfiniBand drivers (the Cisco commercial driver stack and OFED drivers) is not supported; if administrators try to install one of these drivers on top of the other, the system displays a warning message. To help avoid other conflicts between the RPM packages in the Cisco OFED Roll and Red Hat Enterprise Linux, the Cisco RPM packages are installed as part of a post-installation process. A list of conflicting RPM packages is used to back up the native RPM packages from the Red Hat RPMS folder. Other packages that may cause a conflict, such as the Open MPI RPM packages from the Cisco OFED Roll, are backed up to /opt/cisco_ofed/backup/rpm and replaced when the Cisco OFED Roll is uninstalled. A startup script is placed at /etc.profile.d to load the Cisco openmpi_gcc module file as a default environment, as provided by the cisco-modulefiles RPM package. The Cisco firmware bin files are packaged into an RPM package installed in /opt/cisco_ofed/firmware.

High-performance, low-latency HPC clusters

The OpenFabrics Enterprise Distribution is designed to provide a standardized open source software stack for the InfiniBand interconnect, helping simplify the creation and management of HPC clusters. Deploying OFED 1.2 on InfiniBand–based clusters of Dell PowerEdge servers in conjunction with Platform OCS and the Cisco OFED Roll enables organizations to easily build high-performance, low-latency HPC cluster environments.

Munira Hussain *is a systems engineer and adviser in the Dell High-Performance Computing Group. She specializes in HPC systems and architecture design, and her areas of interest include high-speed, low-latency interconnect networks such as InfiniBand and application tuning and benchmarking. Munira has a bachelor's degree in Electrical Engineering from the University of Illinois at Urbana-Champaign.*

Arun Rajan *is a systems engineer in the Dell Scalable Systems Group. His current interests and responsibilities include HPC cluster management, computing packages, performance benchmarking, and product development. He has a B.E. in Electronics and Communications Engineering from the National Institute of Technology, Tiruchirappalli, and an M.S. in Computer and Information Science from the Ohio State University.*

Sreeram Vedantham *is a development engineer adviser in the Dell High-Performance Computing Group. His current interests include high-speed interconnects, HPC cluster stacks, and schedulers.*

Jacob Liberman *is a development engineer in the Dell Scalable Systems Group. Jacob has a master's degree in Instructional Technology from the University of Texas at Austin and holds multiple industry certifications.*

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Hussain, Munira, Rinku Gupta, and Tong Liu. "Using OpenFabrics InfiniBand for HPC Clusters." *Dell Power Solutions*, November 2006. DELL.COM/downloads/global/power/ps4q06-20060430-Hussain.pdf

Rajan, Arun, Tong Liu, Yung-Chin Fang, Garima Kochhar, and Ron Pepper. "Customizing Management of HPC Clusters." *Dell Power Solutions*, February 2007. DELL.COM/downloads/global/power/ps1q07-20060427-Fang.pdf

MULTI-CORE SCALABILITY WITH FLUENT AND MICROSOFT WINDOWS COMPUTE CLUSTER SERVER 2003

BY AZIZ GULBEDEN
AMINA SAIFY

Using the FLUENT computational fluid dynamics application in conjunction with Microsoft® Windows® Compute Cluster Server 2003 enables cluster users to easily run large simulations on high-performance computing clusters. This article describes tests performed to evaluate the scalability of these types of clusters when using multi-core processing to help increase performance.

Related Categories:

Computational fluid dynamics (CFD)

Dell PowerEdge servers

High-performance computing (HPC)

Microsoft Windows Compute Cluster Server 2003

*Visit **DELL.COM/PowerSolutions** for the complete category index.*

In high-performance computing (HPC) cluster environments, using multi-core processors can significantly increase application performance. Although the advantages of additional cores can be significant, bottlenecks from other components can still restrict cluster scalability.

To help demonstrate how cluster performance scales with additional processing cores, in May 2007 Dell engineers tested a cluster based on Microsoft Windows Compute Cluster Server 2003 (CCS) using the FLUENT computational fluid dynamics (CFD) application. The results showed that although quad-core processors did provide higher performance than dual-core processors, the limitations of Gigabit Ethernet networks tended to inhibit scalability when using more than one compute node, primarily because of the increase in communication time when additional servers are involved in the calculation.

Software and hardware in the test environment

The test environment consisted primarily of CCS and the FLUENT CFD application running on a cluster of Dell™ PowerEdge™ servers.

Microsoft Windows Compute Cluster Server 2003

CCS consists of the Microsoft Windows Server® 2003 Compute Cluster Edition OS—a version of Windows Server 2003 designed for HPC cluster environments—and the Compute Cluster Pack (CCP). The CCP adds crucial cluster components, such as a job scheduler allowing cluster users to submit cluster jobs; Microsoft Message Passing Interface (MS-MPI), which provides middleware for parallel jobs; and Compute Cluster Administrator, which provides node management tools.

Windows-based clusters consist of a head node, which enables cluster administration, and compute nodes, which run submitted jobs. Compute node software is typically installed remotely over a network by running Microsoft Remote Installation Services or Windows Deployment Services (WDS) on the head node. In the test environment, the cluster ran Windows Server 2003 Compute Cluster Edition with Service Pack 2 (SP2) and the CCP with SP1.

FLUENT

FLUENT is a commonly used CFD application in HPC cluster environments, designed to model fluid flow and heat transfer in complex geometries. It includes nine CFD application benchmarks for testing FLUENT performance on different hardware. These applications are categorized as small, medium, and large: small problems contain less than 100,000 cells, medium problems contain from 100,000 to 500,000 cells, and large problems contain more than 500,000 cells. FLUENT 6.3, the version used in the test environment, provides support for 64-bit Windows platforms such as CCS.

> *"Although quad-core processors did provide higher performance than dual-core processors, the limitations of Gigabit Ethernet networks tended to inhibit scalability."*

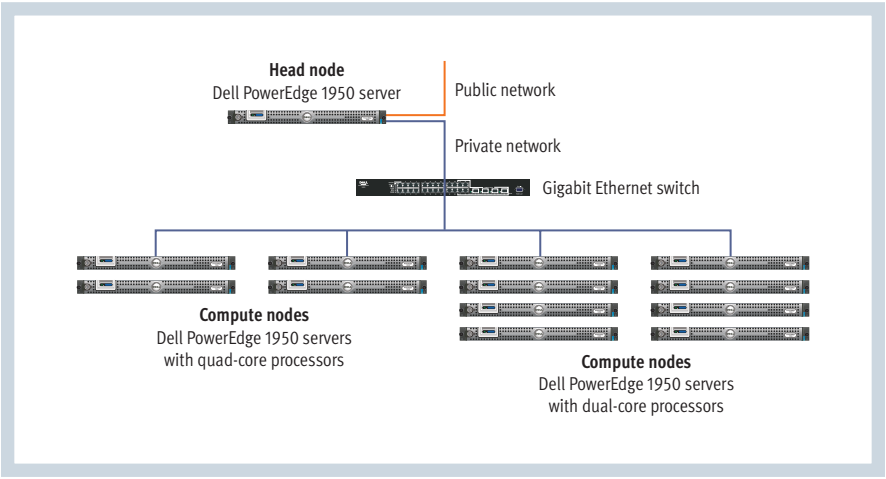


Figure 1. *Cluster architecture used in the test environment*

Cluster architecture

Figure 1 shows the cluster architecture used in the test environment. The cluster consisted of 12 compute nodes: 8 Dell PowerEdge 1950 servers with dual-core Intel® Xeon® 5160 processors at 3.00 GHz and 4 PowerEdge 1950 servers with quad-core Intel Xeon X5355 processors at 2.66 GHz. Each node had 4 GB of RAM and was connected to a Gigabit Ethernet network.

The head node, a PowerEdge 1950 with dual-core Intel Xeon 5150 processors at 2.66 GHz, controlled the Microsoft Active Directory® domain, ran the WDS server that enabled the test team to deploy the compute nodes over the network, and stored the compute node OS image. Because the storage and network drivers were not included in the OS by default, the test team added them to the image manually.[1]

On the head node, the test team created a shared folder that allowed the compute nodes to access the FLUENT application on the head node, and also stored output and error messages so that the team could troubleshoot if a job failed to run. The head node also hosted FLUENT Software License Manager, which restricts the number of processors used by the application. To launch the jobs, the test team used FLUENT Launcher, which includes an option to use the CCS job scheduler to select nodes for running the applications.

Cluster users can check job status using the Compute Cluster Job Manager that comes with the CCP, which displays errors encountered during execution and allows them to cancel or resubmit jobs if needed.

Test results: Performance and scalability with multi-core processors

The engineering team ran small, medium, and large FLUENT benchmarks and measured performance by a rating denoting the number of benchmarks that a server could run over a 24-hour period. To help demonstrate how additional cores can affect performance and scalability, these ratings were normalized to show relative performance compared with a single process.

Figure 2 compares the relative performance rating of a single compute node with two dual-core processors with that of a single compute node with two quad-core processors. These servers ran four and eight processes, respectively, for a variety of FLUENT problem sizes. The server with quad-core processors

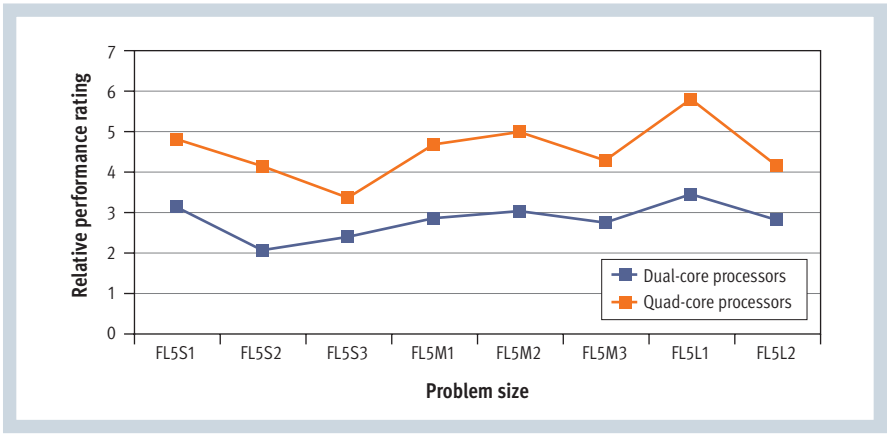


Figure 2. *Relative performance ratings of a single server with two dual-core processors and a single server with two quad-core processors*

[1] For more information, see "Deploying Microsoft Windows Compute Cluster Server 2003 on Dell PowerEdge Servers," by Ron Pepper and Victor Mashayekhi, Ph.D., in *Dell Power Solutions*, November 2006, DELL.COM/downloads/global/power/ps4q06-20070153-Pepper.pdf.

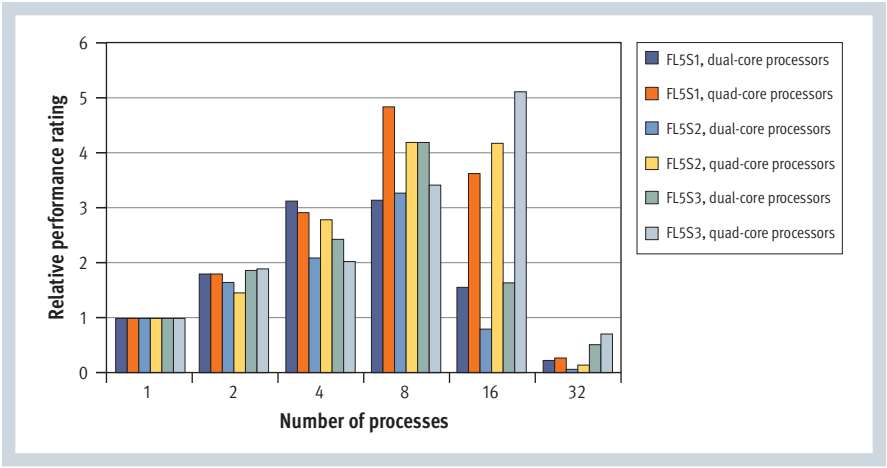


Figure 3. *Relative performance ratings for small FLUENT problem sizes*

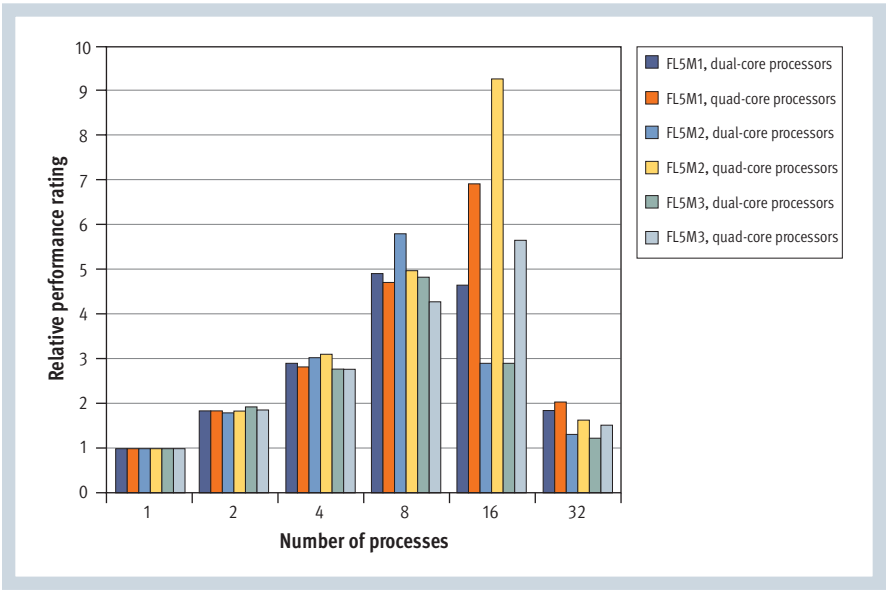


Figure 4. *Relative performance ratings for medium FLUENT problem sizes*

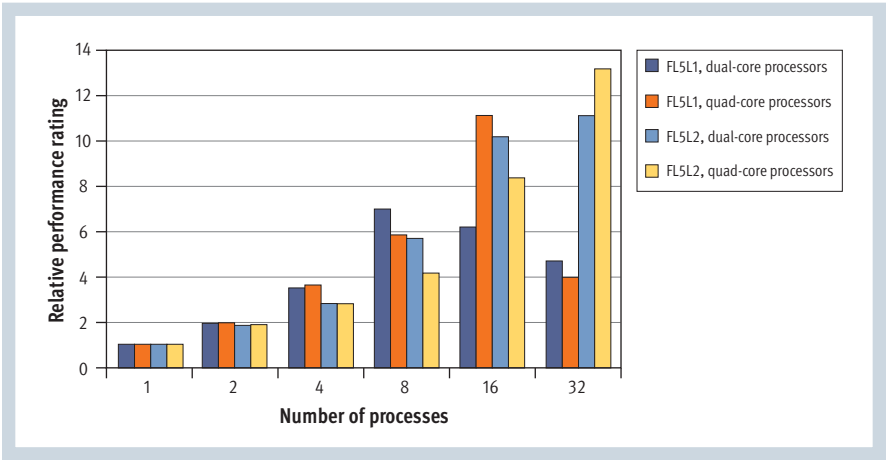


Figure 5. *Relative performance ratings for large FLUENT problem sizes*

outperformed the server with dual-core processors for all problem sizes tested, even though the quad-core processors had a lower clock speed than the dual-core processors.

Figures 3–5 show the relative performance ratings for a variety of small, medium, and large problem sizes. As these figures show, servers with both dual- and quad-core processors have difficulty scaling beyond eight processes (two servers with dual-core processors, or one server with quad-core processors), a result attributable to the performance limitations of the Gigabit Ethernet network. The most notable exception was the FL5L2 problem size, in which the amount of computation on each node is greater than the communication required between nodes.

Cluster scalability with multi-core processing

As the test results showed, a single server with quad-core processors can significantly outperform a single server with dual-core processors when running the FLUENT application in an HPC cluster based on Microsoft Windows Compute Cluster Server 2003. However, the limitations of the Gigabit Ethernet network in the test environment prevented the cluster from scaling well when connecting additional nodes to handle additional processes.

Aziz Gulbeden *is a systems engineer in the Scalable Systems Group at Dell. His current areas of focus include Microsoft Windows–based compute clusters and scalable file and storage systems. He has a B.S. in Computer Engineering from Bilkent University and an M.S. in Computer Science from the University of California, Santa Barbara.*

Amina Saify *is a member of the Scalable Systems Group at Dell. Her current areas of focus include cluster management, cluster file systems, storage, and performance benchmarking. Amina has a bachelor's degree in Computer Science from Devi Ahilya University and a master's degree in Computer and Information Science from the Ohio State University.*

EVALUATING THE MYRINET-10G INTERCONNECT ON DELL POWEREDGE SERVERS

BY SREERAM VEDANTHAM
SHIVARAJ NIDONI
MUNIRA HUSSAIN

The Myricom Myrinet-10G (Myri-10G) interconnect can offer high bandwidth and low latency in high-performance computing environments. This article describes the Myri-10G interconnect architecture and software stack as well as test results comparing the latency, throughput, and performance of Myri-10G, Myrinet-2000, and Gigabit Ethernet on ninth-generation Dell™ PowerEdge™ servers.

Related Categories:

Dell PowerEdge servers

High-performance computing (HPC)

Interconnects

Myrinet

Visit ***DELL.COM/PowerSolutions*** *for the complete category index.*

When building a high-performance computing (HPC) system for communication-intensive applications, selecting an interconnect that offers high bandwidth and low latency at a good price/performance ratio can be critical. The Myricom Myrinet-10G (Myri-10G) interconnect can do just that, binding multiple nodes and enabling intra-node cluster communication to form the backbone of an HPC architecture.

Introducing Myri-10G

Myri-10G, based on 10 Gigabit Ethernet (10GbE) technology, is the fourth-generation Myrinet interconnect, following the third-generation Myrinet-2000. Myri-10G is supported in two modes: Ethernet mode and Myrinet Express (MX) mode. The MX protocol offers kernel bypass with high-bandwidth, low-latency software support for intra-process cluster communication.

Myri-10G network interface cards (NICs) can be connected to Myri-10G or 10GBase-T switches. The current Myri-10G switch is a 14U enclosure with a 128-port switch that supports both 10GBase-CX4 copper interfaces and 10G-Q quad ribbon-fiber cables for 10GbE Attachment Unit Interface (XAUI).

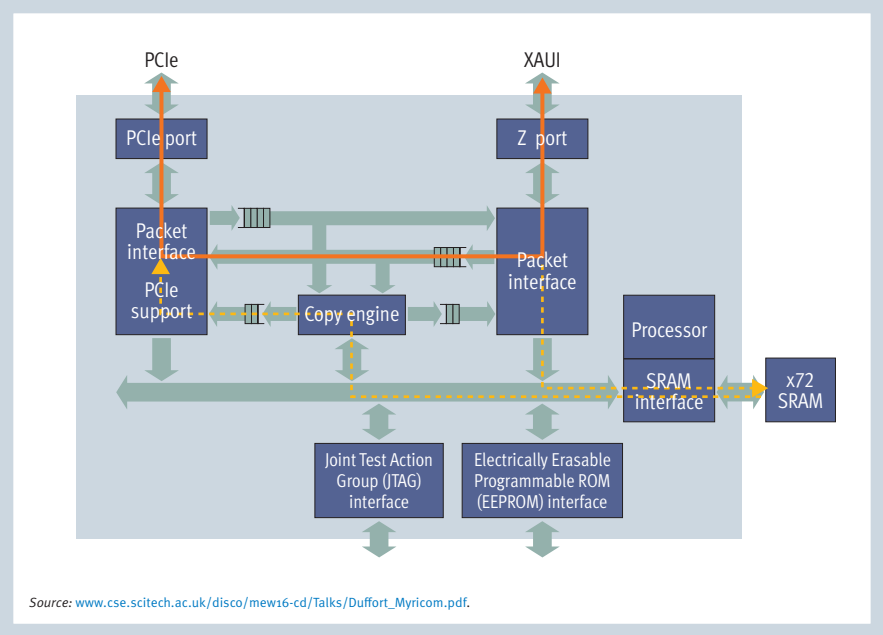
The x8 PCI Express (PCIe) bandwidth supports 2+2 GB/sec full duplex I/O transfer capacity, which is wide enough to accommodate Myri-10G NICs that support 1.25+1.25 GB/sec bidirectional bandwidth. The NIC offers a total data rate of 10+10 Gbps full duplex. (Myrinet-2000, by comparison, offers a total bandwidth of 2+2 Gbps.)

At the physical layer, the Myri-10G NIC is available in two types: 10GBase-CX4, which connects to a switch through copper cables, and 10GBase-R, which connects to a switch through fiber cables. Both types can operate in Ethernet mode or MX mode; in MX mode, they can deliver low-latency, high-bandwidth kernel bypass Message Passing Interface (MPI) communication that helps reduce host processor utilization.

Myri-10G hardware architecture

The Myri-10G hardware architecture consists of processors, memory, and firmware. This architecture enables the offloading of network protocol processing, minimizing host processor utilization and enabling direct low-latency kernel bypass communication with applications. The Myri-10G NIC has a very-large-scale integration (VLSI) chipset known as the Lanai Z8E, and includes an internal processor at 313 MHz and built-in static RAM (SRAM).

The architecture also includes internal packet buffers. One advantage of on-chip buffers is that they provide most of the required memory bandwidth needed during data communication, thus limiting SRAM memory use on the NIC. The NIC has approximately 2 MB of local SRAM available with a bandwidth of 2,400 MB/sec, which is primarily used for firmware execution and packet header buffering. Packets are transferred through dynamic memory allocation, helping reduce host processor utilization.



Source: www.cse.scitech.ac.uk/disco/mew16-cd/Talks/Duffort_Myricom.pdf.

Figure 1. *Chipset architecture on Myri-10G NICs*

Figure 1 shows how packets are transferred and handled within the Myri-10G NIC. The yellow dotted line shows how packets are passed to and from the internal buffer memory on a single NIC; the orange line represents the physical contact connections of the NIC to the x8 PCIe slot in an HPC node. XAUI can be converted to either 10GBase-CX4 copper or 10GBase-R fiber. Administrators can choose the connection type based on cost and the distance between the nodes and the switch.

Myri-10G software protocols

Myri-10G NICs support both Gigabit Ethernet and MX. MX is a low-level MPI implementation designed to provide low latency for small messages, support a virtually unlimited number of pending send and receive requests, and handle configurations with no memory registration.

The type of switch to which the NIC is connected determines whether network packets are transferred using TCP/IP or MX. In a cluster configuration with the default 10GbE drivers installed, the NIC uses TCP/IP; if MX drivers are installed, however, the NIC uses MX.

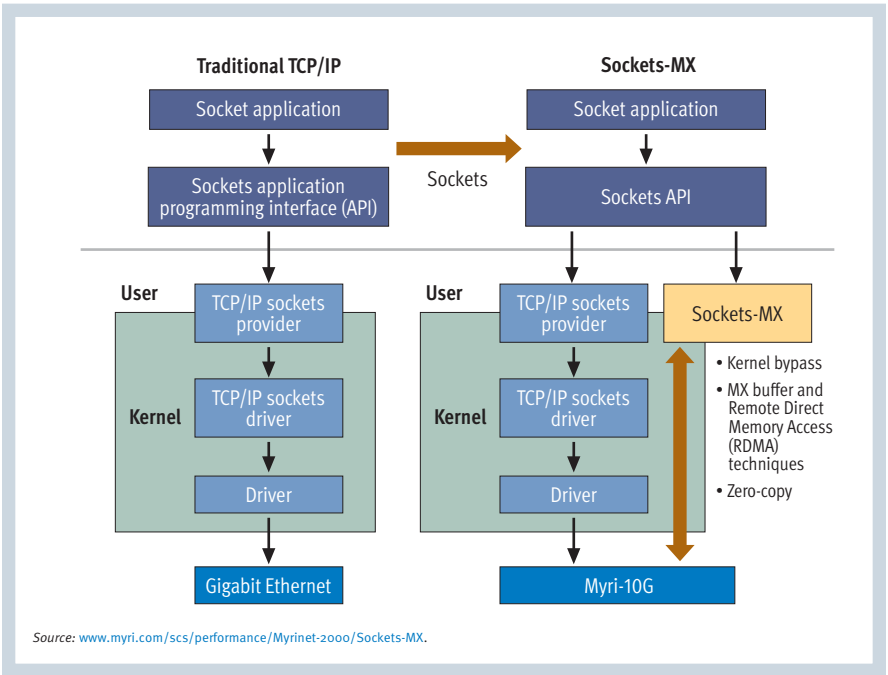
When MX drivers are installed, if the NIC is connected to a 10GbE switch, then MX uses its TCP/IP mode. If the NIC is connected to a Myri-10G switch—which is designed to handle MX packets—then MX uses Sockets-MX (see Figure 2). Unlike TCP/IP, Sockets-MX bypasses the traditional kernel stack, which can result in the MX protocol having lower latency than TCP/IP, in which the kernel traps typically take up communication time and create overhead.

Testing Myri-10G on ninth-generation Dell PowerEdge servers

In May 2007, Dell engineers from the High-Performance Computing Group performed tests to compare the performance of Myri-10G in MX mode with that of Myrinet-2000 in GM mode and Gigabit Ethernet. The Dell team used the Ohio State University (OSU) MPI-level latency test to evaluate latency, the Intel MPI Benchmarks (IMB) suite to evaluate unidirectional and bidirectional bandwidth, and High-Performance Linpack (HPL) to evaluate performance in clusters using each of the three interconnects.[1] They compiled the tests to run on Myri-10G using the MPICH MPI implementation.

Cluster configurations in the test environment

Figure 3 details the two cluster configurations used in the test environment, which were based on Dell PowerEdge servers using either Intel® Xeon® processors or AMD Opteron™ processors.



Source: www.myri.com/scs/performance/Myrinet-2000/Sockets-MX.

Figure 2. *Traditional TCP/IP model compared with Myri-10G Sockets-MX*

[1] For more information on these benchmarks, visit mvapich.cse.ohio-state.edu/benchmarks, www.intel.com/cd/software/products/asmo-na/eng/307696.htm#mpibenchmarks, and www.netlib.org/benchmark/hpl.

The team ran all tests on both clusters, then chose the best performance for the final results.

The BIOS versions on the cluster nodes support write-combining. Write-combining is supported on Myri-10G NICs by default on the PCIe chipset, and helps optimize Myri-10 performance by allowing data to be combined and temporarily stored in the buffer, then written (in "write-bursts") whenever the buffer is full. This approach can be particularly helpful in frame buffers, and can increase the performance of write operations, especially when attempting to avoid excess switching overhead between read and write operations.

Test results: Latency, bandwidth, and performance

Figure 4 shows normalized latency for Myri-10G and Myrinet-2000 as a percentage of Gigabit Ethernet latency, as measured using the OSU latency test. Myri-10G had a very low latency of 5.4 percent of the total latency of Gigabit Ethernet, while Myrinet-2000 had approximately 12 percent of the total latency of Gigabit Ethernet.

Figure 5 shows normalized unidirectional bandwidth for Myri-10G and Myrinet-2000 as a multiple of Gigabit Ethernet unidirectional bandwidth, as measured using the IMB PingPong test. The test team measured Myri-10G both with and without the MX_RCACHE=1 variable. When this variable is enabled, MX uses registration cache rather than fetching each communication from memory, helping maximize bandwidth performance with the MX driver stack. Myri-10G with this variable provided 12.11 times the bandwidth of Gigabit Ethernet, or approximately 91 percent of the theoretical unidirectional bandwidth offered by the NIC. Myri-10G without this variable provided 9.97 times the bandwidth of Gigabit Ethernet, or approximately 75 percent of the theoretical unidirectional bandwidth offered by the NIC.

Figure 6 shows normalized bidirectional bandwidth for Myri-10G and Myrinet-2000 as a multiple of Gigabit Ethernet bidirectional bandwidth, measured using the IMB SendRecv test. In these tests, Myri-10G with the MX_RCACHE=1 variable provided 21.4 times the bandwidth of Gigabit Ethernet, while Myri-10G without this variable provided 17.8 times the bandwidth of

	Intel Xeon processor–based cluster	AMD Opteron processor–based cluster
Servers	Four Dell PowerEdge 1950 servers	Four Dell PowerEdge SC1435 servers
Processors	Two dual-core Intel Xeon X5355 processors at 2.66 GHz with a 1,333 MHz frontside bus and 4 MB L2 cache per socket	Two dual-core AMD Opteron 2218 processors at 2.6 GHz with a 1 MB L2 cache per socket
Memory	4 GB, using fully buffered dual in-line memory modules (DIMMs)	8 GB, using double data rate 2 (DDR2) DIMMs
BIOS version	1.3.3	1.1.2
Riser cards	x8 PCIe	
Switches	Myri-10G, comprising 10G-SW16LC-8C line cards with eight 10GBase-CX4 front panels	
Cable type	10GBase-CX4	
OS	Platform OCS 4.4.0	

Figure 3. *Cluster configurations used in the test environment*

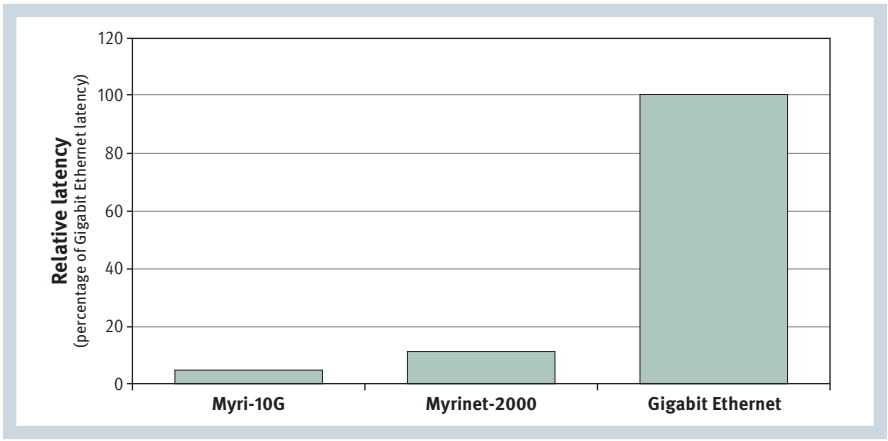


Figure 4. *Myri-10G and Myrinet-2000 latency as a percentage of Gigabit Ethernet latency, measured using the Ohio State University latency test*

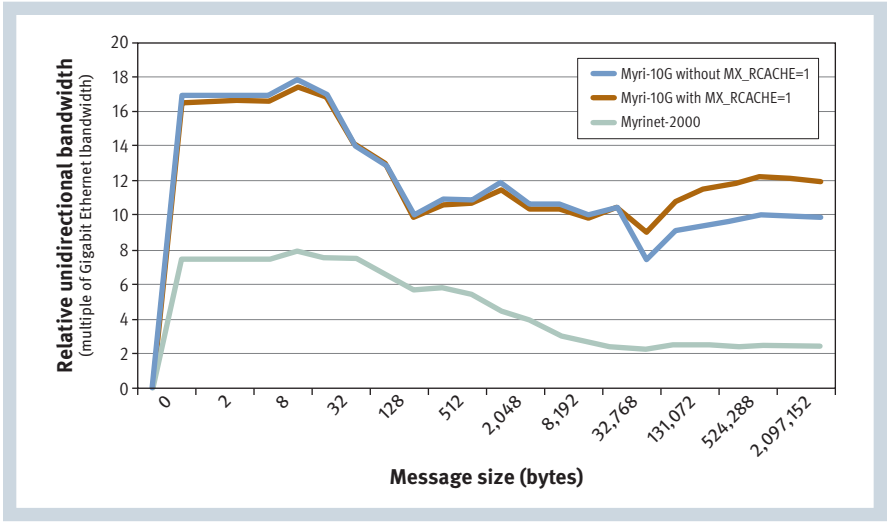


Figure 5. *Myri-10G and Myrinet-2000 unidirectional bandwidth as a multiple of Gigabit Ethernet bandwidth, measured using the Intel MPI Benchmarks PingPong test*

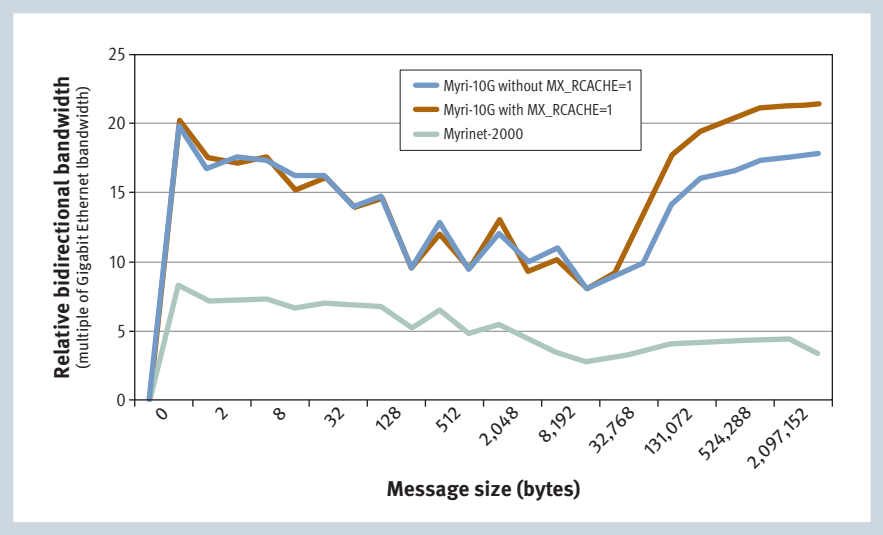


Figure 6. *Myri-10G and Myrinet-2000 bidirectional bandwidth as a multiple of Gigabit Ethernet bandwidth, measured using the Intel MPI Benchmarks SendRecv test*

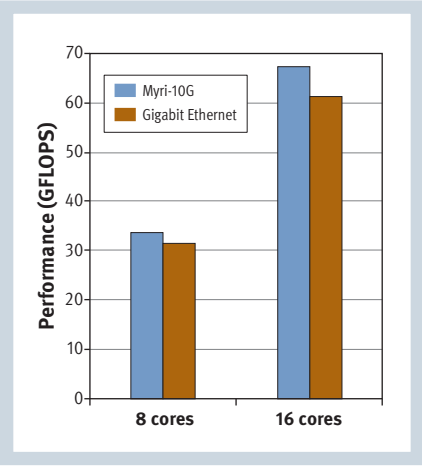


Figure 7. *Myri-10G and Gigabit Ethernet performance, measured using the High-Performance Linpack application*

Gigabit Ethernet. As with unidirectional bandwidth, Myri-10G with and without this variable provided approximately 91 percent and 75 percent of the theoretical bidirectional bandwidth offered by the NIC, respectively.

Figure 7 compares the performance in gigaflops of Myri-10G and Gigabit Ethernet, as evaluated when running HPL, a communication- and computation-intensive application well suited to serve as a cluster-level benchmark. The difference between the two interconnects is emphasized when scaling up the clusters. Using Myri-10G with 8 cores (two nodes) provided 81 percent of the processors' theoretical maximum gigaflops, as compared with 75.48 percent when using Gigabit Ethernet. When scaled to 16 cores (four nodes), Myri-10G provided 80.84 percent of the theoretical maximum gigaflops, as compared with 73.73 percent when using Gigabit Ethernet. As the number of processing cores increased, communication across nodes became a bottleneck for Gigabit Ethernet, reducing performance relative to the theoretical maximum performance possible for a given number of cores.

Deploying Myri-10G on ninth-generation Dell PowerEdge servers

Dell has partnered with Myricom to deliver a pre-validated cluster solution using Myri-10G NICs and the 128-port Myri-10G switch in a 14U enclosure. Dell servers support these NICs and the switch in MX mode only. The switch enclosure houses the hot-swappable power supplies, cooling and network monitoring systems, a status display, and a backplane that supports up to 21 hot-swappable line cards. It also includes a monitoring line card that fits into the left line-card slot. The four center slots also contain line cards, with eight 16-port crossbar switches that can make up the spine of a Clos network; Dell-supported 10GBase-CX4 10G-SW16LC-8C line cards can be used in the 16 remaining slots.

Dell recommends deploying clusters using Platform Open Cluster Stack (OCS), which is pre-validated as a part of Dell HPC clusters. The Myri-10G software stack and drivers form a separate OCS Roll that administrators can first install on the front end during initial configuration and then deploy on the compute nodes through Preboot Execution Environment (PXE) kickstart images. Device drivers are loaded in the kernel space during compute node installation. Platform OCS enables both Myri-10G and Myrinet-2000 Rolls to coexist; these Rolls are different from and independent of each other.

Boosting HPC cluster performance with Myri-10G

As shown in Dell benchmarking tests, the Myricom Myri-10G interconnect can provide low latency and increased throughput in communication-intensive environments, including those with parallel applications running different MPI stacks such as MPICH and Open MPI. Myri-10G also remains interoperable with 10GbE and can minimize host processor utilization when running using MX. Enterprises can take advantage of these features to increase the performance of their HPC applications.

Sreeram Vedantham *is a development engineer adviser in the Dell High-Performance Computing Group. His current interests include high-speed interconnects, HPC cluster stacks, and schedulers.*

Shivaraj Nidoni *is a senior engineering analyst in the Dell High-Performance Computing Group. His current interests include high-end interconnects, parallel file systems, and HPC clustering packages. Shivaraj has an M.Tech. in Computer Network Engineering from Visvesvaraya Technological University.*

Munira Hussain *is a systems engineer and adviser in the Dell High-Performance Computing Group specializing in HPC systems and architecture design. She has a bachelor's degree in Electrical Engineering from the University of Illinois at Urbana-Champaign.*

INTEL CLUSTER READY AND PLATFORM OPEN CLUSTER STACK: CLUSTERS MADE SIMPLE

The Intel® Cluster Ready program is designed to provide a common standard for high-performance computing (HPC) clusters, helping organizations design and build seamless, compatible configurations. Integrating the standards and tools provided by this program with Platform™ Open Cluster Stack and certified Dell™ clusters can help significantly simplify the deployment and management of HPC clusters.

Related Categories:

High-performance computing (HPC)

Parallel systems

Platform Computing

Visit ***DELL.COM/PowerSolutions*** *for the complete category index.*

As the cost of high-performance computing (HPC) falls, a growing range of problems have become economical to solve with compute clusters based on commodity hardware components. These problems range from risk management for insurance portfolios to design optimization and durability studies on automobile and aerospace components, which can now be solved with levels of compute power formerly reserved for only the biggest and most costly problems. Even in areas where non-cluster HPC systems have been in use for some time, such as oil and gas reservoir simulation and life sciences clinical studies, commodity hardware enables analysis of larger problem sizes with more fidelity than was previously economically feasible.

According to IDC, the HPC technical server market, which is 50 percent compute clusters, is growing dramatically in the divisional, departmental, and workgroup areas. This growth is largely attributable to the use of commodity components, which have dramatically reduced the price/performance ratio of these systems. However, IDC still expresses concern that adoption may be limited, because "clusters are increasingly complex to deploy and manage" and require advanced or specialist skill sets for IT personnel.[1]

The source of this complexity is the cluster architecture itself. Cluster servers are fundamentally different from traditional symmetric multiprocessing servers, because they comprise individual units for processing and storage. In addition, using clusters for either multiple jobs or to run software applications that can use multiple processors simultaneously requires high-speed interconnects as well as workload management middleware. Because of these differences, clusters also require different approaches to specification, installation, and management than are used in traditional IT environments.

Until recently, it was difficult to ensure that HPC clusters met a minimum set of standards: each cluster may have had different hardware and software components, and the resulting combinations may or may not have functioned in the same way. To help avoid this problem, Intel has collaboratively developed the Intel Cluster Ready program and technology package with original equipment manufacturers, channel members, and independent software vendors (ISVs). By taking advantage of the standards and tools provided by this program, and combining them with Platform Open Cluster Stack (OCS) software and certified Dell HPC

This product includes software developed by the Rocks™ Cluster Group at the San Diego Supercomputer Center at the University of California, San Diego, and its contributors.

[1] "Intel Cluster Ready," by IDC, Doc #207312, June 2007.

"Platform OCS helps simplify the deployment and management of Intel Cluster Ready–certified clusters by installing and configuring Intel Cluster Ready software components on Dell HPC platforms."

clusters, organizations can help significantly simplify the deployment and management of HPC clusters.

Introducing Intel Cluster Ready and Intel Cluster Checker

Intel and its partners have created the Intel Cluster Ready program to help simplify the definition, acquisition, installation, and management of HPC clusters for organizations without prior experience in cluster computing and those working to increase their technical computing capacity. By incorporating certified hardware, cluster system software, application software, and cluster-ready configurations, this program helps reduce both deployment time and total cost of ownership—both of which can be critical in environments where HPC applications are delivering essential competitive and strategic advantages.

Intel Cluster Ready provides a reference specification for ISVs and system builders to help validate HPC clusters as well as a set of configurations describing in detail how to combine components to create an Intel Cluster Ready–certified cluster. For IT organizations, the key feature of this program is that it specifies a common basis for clusters, allowing them to select from a variety of hardware and software components based on their cluster's purpose and helping ensure that ISV applications that work on one certified cluster can also run reliably on a different certified cluster. This common basis significantly simplifies the processes of designing, building, acquiring, and deploying clusters based on Intel components, increasing flexibility and helping reduce total cost of ownership.

The Intel Cluster Ready specification is a key part of the program, but the program consists of more than just documentation. The Intel Cluster Checker, a script-based tool that performs direct computational tests and measurements, helps both vendors and IT organizations ensure conformance to the specification, provides an objective measure of system performance, and can assist in troubleshooting. Figure 1 illustrates the architecture of the Intel Cluster Checker engine. This tool is designed to significantly reduce deployment time while increasing uptime for certified clusters.

If a cluster passes all of the Cluster Checker tests, it is considered Intel Cluster Ready certified. Organizations can also use this tool to help ensure that the cluster continues operating properly and within the specification simply by running tests periodically on the cluster and comparing the results with those of previous tests. Doing so helps detect deviations from the original cluster certification to help ensure that the cluster remains certified.

Combining Platform OCS and certified Dell HPC clusters

Platform OCS is a pre-integrated, vendor-certified, modular software stack designed to streamline the deployment and management of clusters running the Linux® OS. Backed by available global 24/7 enterprise support, it transparently integrates open source and commercial software into a single consistent cluster operating environment. Platform OCS helps simplify the deployment and management of Intel Cluster Ready–certified clusters by installing and configuring Intel Cluster Ready software components on Dell HPC platforms (see Figure 2).

Platform Load Sharing Facility (LSF®) HPC, a powerful, comprehensive, policy-driven workload management application for engineering and scientific distributed computing environments, works in conjunction with Platform OCS to intelligently schedule parallel and serial workloads, helping maximize available computing resources. By utilizing hardware-specific integrations, Platform LSF HPC and Platform

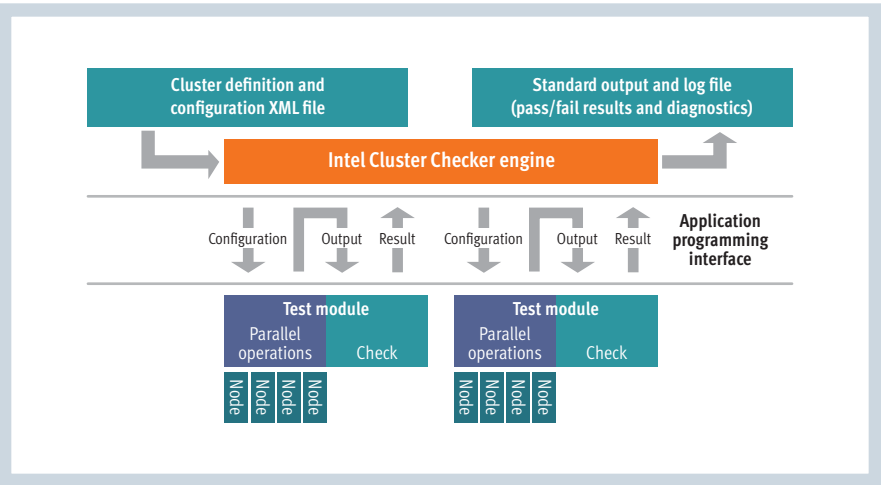


Figure 1. *Intel Cluster Checker engine architecture*

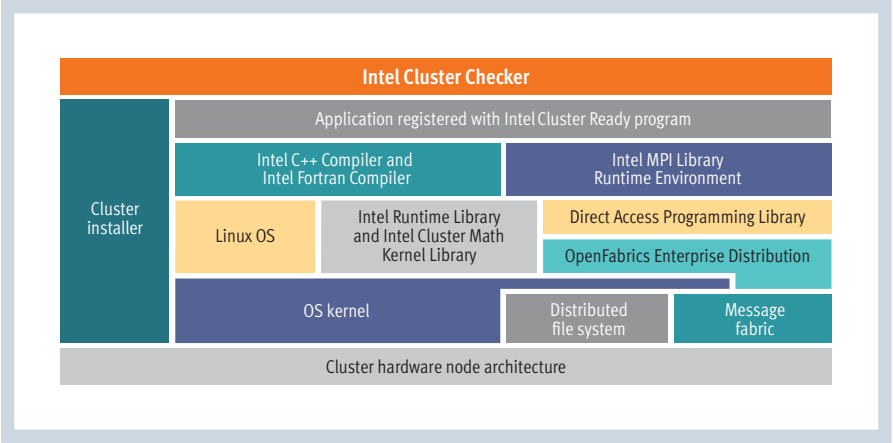


Figure 2. *Intel Cluster Ready components installed and configured by Platform OCS*

OCS also enable organizations to take advantage of the high-performance network interconnects available on clustered systems and supercomputers.

Combining Platform software with Dell hardware allows organizations to easily create and deploy HPC clusters using industry-standard components. HPC clusters incorporate a diverse array of hardware components, and the appropriate choice of computers, processors, memory, hard drives, storage arrays, network devices, cabling, switches, and power supplies depends on the cluster's purpose. Organizations must carefully match the hardware with their goal to help achieve the desired performance. Hardware from Dell—a leader in HPC—can provide a standard, reliable reference platform when building clusters.

Because Platform OCS is already a part of several Intel Cluster Ready configurations, using it helps alleviate the need for organizations to assemble, configure, and install the necessary components either manually or by using other cluster management tools. They can use Platform OCS to quickly install and configure Intel Cluster Ready–certified Dell HPC clusters with the following components:

- Certified Dell hardware
- A Message Passing Interface (MPI) implementation such as Open MPI or Intel MPI Library
- The Intel Runtime Library, including the Intel MPI Library Runtime Environment
- The OpenFabrics Enterprise Distribution stack (optional)

For example, an Intel Cluster Ready configuration might include Platform OCS with the Intel Cluster Checker Roll as well as the following hardware components:

- One Dell PowerEdge™ 2950 server as the front-end node
- 12 Dell PowerEdge 1950 servers as the compute nodes
- One 16-port Dell PowerConnect™ switch
- KVM (keyboard, video, mouse) over IP switch, cables, server rack, cable management system, and power distribution system

Building a standard for simplified cluster deployment

The Intel Cluster Ready program is designed to let organizations easily deploy and manage HPC clusters, helping eliminate the need to create custom implementations in which they must individually install and configure each application while modifying the cluster hardware and system software to meet these applications' requirements. This program also helps significantly simplify the work required by commercial and noncommercial application vendors, who can focus on certifying their applications for Intel Cluster Ready configurations rather than on porting and configuring the applications for different potential combinations of hardware and software. By integrating the Platform OCS software stack with Intel Cluster Ready–certified Dell HPC clusters, organizations can easily create seamless cluster environments to help meet their HPC requirements.

"Combining Platform software with Dell hardware allows organizations to easily create and deploy HPC clusters using industry-standard components."

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Intel Cluster Ready:
www.intel.com/go/cluster

Platform OCS:
www.platform.com/Products/Platform.OCS

Dell HPC cluster solutions:
DELL.COM/HPCC

This product includes software developed by the Rocks™ Cluster Group at the San Diego Supercomputer Center at the University of California, San Diego, and its contributors.

Extending OpenMP to Clusters

OpenMP is a well-known parallel programming paradigm for shared-memory multiprocessors. In the past, OpenMP has been confined to Symmetric Multi-Processing (SMP) machines and teamed with Message Passing Interface (MPI) technology to make use of multiple SMP systems. A new system, Cluster OpenMP, is an implementation of OpenMP that can make use of multiple SMP machines without resorting to MPI. This advance has the advantage of eliminating the need to write explicit messaging code, as well as not mixing programming paradigms. The shared memory in Cluster OpenMP is maintained across all machines through a distributed shared-memory subsystem. Cluster OpenMP is based on the relaxed memory consistency of OpenMP, allowing shared variables to be made consistent only when absolutely necessary.

Performance Considerations for Cluster OpenMP

Some memory operations are much more expensive than others. To achieve good performance with Cluster OpenMP, the number of accesses to unprotected pages must be as high as possible, relative to the number of accesses to protected pages. This means that once a page is brought up-to-date on a given node, a large number of accesses should be made to it before the next synchronization. In order to accomplish this, a program should have as little synchronization as possible, and re-use the data on a given page as much as possible. This translates to avoiding fine-grained synchronization, such as atomic constructs or locks, and having high data locality.

The OpenMP memory model allows individual reads and writes to memory to be done in any order, as long as the synchronization operations (flushes) are done in a strict order–the same order in which they appear in the original user's program. The lack of ordering between reads and writes to memory makes their concurrent execution possible, but all flushes in a program must be serialized, adding overhead to the program.

Latency to L1	**1-2 cycles**
Latency to L2	**5-7 cycles**
Latency to L3	**12-21 cycles**
Latency to memory	**180-225 cycles**
Gigabit Ethernet latency to remote node	**~ 28000 cycles**
InfiniBand* latency to remote node	**~ 23000 cycles**

Figure 1. Itanium® processor latency to cache and memory compared with messaging latency to remote nodes. Cluster OpenMP currently runs on Itanium®-based platforms and on systems based on processors that support Intel® EM64T running Linux*.

Figure 1 shows the number of processor cycles required for an access to different levels of cache and the latency to access a value in the memory of a remote node. This shows that access to the memory of a remote node is approximately 100 times slower than access to the local memory, and thousands of times slower than access to a value in cache. This comparison should drive home the point that local, rather than remote, memory should be used as much as possible.

Cluster OpenMP provides shared memory across a cluster for an OpenMP program. It takes advantage of the relaxed memory model of OpenMP to optimize the memory accesses in an OpenMP program.

Cluster OpenMP does not perform well for all types of programs, but programs with certain characteristics can achieve reasonably good performance on a cluster, compared with attainable performance on a hardware shared memory machine.

This white paper can be downloaded in its entirety from http://assets.devx.com/goparallel/19403.pdf and shows performance results for a set of applications, showing that most were able to achieve greater than 70 percent of the performance of OpenMP programs run on a shared memory machine.

"YOUR COMPETITIVE EDGE IS BECOMING DULL"

-CIO MAGAZINE

IT COMPLEXITY INHIBITS GROWTH

But Dell is ready to change everything with a promise to re-invent information technology as you know it. Simply put, we'll cut the cost and complexity of IT once and for all. Learn more at www.dell.com/simplify.

SIMPLIFY YOUR IT AT DELL.COM/Simplify