DELL POWER SOLUTIONS · MAY 2005

MAY 2005 · $12.95

POWER SOLUTIONS

THE MAGAZINE FOR DIRECT ENTERPRISE SOLUTIONS

Advancing Your Enterprise Storage Strategy

Fast, flexible ways to grow storage platforms cost-effectively

Inside This Issue:

Planning for 4 Gbps Fibre Channel Storage

Driving Databases with the New Dell PowerEdge 6800 and PowerEdge 6850 Servers

Exploring Microsoft Windows Server 2003 x64 Editions

Advancing Your Enterprise Storage Strategy

Gigabit Ethernet
Multi-Gigabit Ethernet

DELL POWER SOLUTIONS
THE MAGAZINE FOR DIRECT ENTERPRISE SOLUTIONS
MAY 2005

EDITOR'S COMMENTS

STORAGE

MICROSOFT WINDOWS TECHNOLOGY

Dell's scalable enterprise strategy is more flexible than ever. Key storage components now include a budget-conscious iSCSI-based array along with a lineup of outstanding CX series Fibre Channel platforms that are designed to deliver high availability, performance, and capacity for demanding enterprise applications.

Dell/EMC AX100i storage array

TABLE OF CONTENTS

Dell PowerEdge 6850 rack server

SCALABLE ENTERPRISE

NETWORK AND COMMUNICATIONS

SYSTEMS MANAGEMENT

DELL™ POWER SOLUTIONS | THE MAGAZINE FOR DIRECT ENTERPRISE SOLUTIONS

WWW.DELL.COM/POWERSOLUTIONS

EDITORIAL

EDITOR-IN-CHIEF | Tom Kolnowski
MANAGING EDITOR | Debra McDonald
SENIOR EDITORS | Liza Graffeo, Vicki Van Ausdall
CONTRIBUTING AUTHORS | Max A. Benham; Robert Bernstein; Matthew Brisse; Michael Brundridge; Scott M. Callaway; Onur Celebioglu; Bill Dawkins, Ph.D.; Randy DeRoeck; Tracy D. Edwards; Yung-Chin Fang; Jeff Ferris; Ryan Franks; John Fruehe; Brian Greene; Kevin Guinn; Rinku Gupta; Stephen M. Hahn; Landon Hale; Bill Henderson; Bruce Holmes; Jenwei Hsieh, Ph.D.; Saeed Iqbal, Ph.D.; Dave Jaffe, Ph.D.; Harish Jayakumar; Monica Kashyap; Garima Kochhar; Min-John Lee; Ed Matthews; Arrian Mehis; Todd Mitchell; Gajanan Mudaliar; Chandra S. Mugunda; Todd Muirhead; Brian Nelson; Dat Nguyen; Nam Nguyen; Marcos Palacios, Ph.D.; Weimin Pan; Ranjith Purush; Ryan Putman; Paul Rad; Ashish Ray; Scott Rosen; Amina Saify; Ananda Sankaran; Spencer Sells; Sandhya Senapathi; Michael Smith; Hector Valenzuela; Vicki Van Ausdall; Bryant Vo; Chip Webb; and Haihong Zhuo

ART

ART DIRECTOR | Mark Mastroianni
DESIGNER AND ILLUSTRATOR | Cynthia Webb
COVER DESIGN | Mark Mastroianni
CONTRIBUTING PHOTOGRAPHER | Joey Pena/Dell Photo Studio
CONTRIBUTING ARTIST | Kim Wyatt

ONLINE

WEB PRODUCTION | Brad Klenzendorf

SUBSCRIPTION SERVICES

CIRCULATION | Becky James

Subscriptions are free to qualified readers who complete the online subscription form or the subscription reply form on page 89 in this issue. To sign up as a new subscriber, renew an existing subscription, change your address, or cancel your subscription, submit the online subscription form at **www.dell.com/powersolutions_subscribe**, return the subscription reply form by surface mail, or fax the subscription reply form to +1 512.283.0363. For subscription services, please e-mail **us_power_solutions@dell.com**.

ABOUT DELL

Dell Inc., headquartered in Round Rock, Texas, near Austin, is the world's leading direct computer systems company. Dell is one of the fastest growing among all major computer systems companies worldwide, with approximately 47,800 employees around the globe. Dell uses the direct business model to sell its high-performance computer systems, workstations, and storage products to all types of enterprises. For more information, please visit our Web site at **www.dell.com**.

Dell Power Solutions is published quarterly by the Dell Product Group, Dell Inc. Dell Power Solutions, Mailstop 8456, Dell Inc., One Dell Way, Round Rock, TX 78682, U.S.A. This publication is also available online at **www.dell.com/powersolutions**. 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.

ON THE FRONT COVER | Managing racks of Dell/EMC storage

May 2005

DELL™

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.

THIS CHANGES EVERYTHING.

DELL OPENMANAGE™/ALTIRIS®
MANAGEMENT SUITE for DELL SERVERS

OS, Application and Hardware Management. Just one console.

Now, Dell™ PowerEdge™ administrators only need one console to deploy, manage, monitor, patch and update software and hardware for Microsoft® Windows® and Red Hat® Linux® environments. With all those features fully integrated, Dell OpenManage with Altiris Management Suite for Dell Servers helps get systems up and running fast, saving IT time and resources.

Take some time to see for yourself.

Visit dell.com/altiris6 today for a demonstration and whitepaper.

altiris®

GET MORE OUT OF CHANGE. GET MORE OUT OF NOW. DELL

Click www.dell.com/altiris6 Call 1.866.212.9344
toll free

Dell cannot be responsible for errors in typography or photography. Dell, PowerEdge, Dell OpenManage and the Dell logo are trademarks of Dell Inc. Altiris is a registered trademark and Server Management Suite is a trademark of Altiris, Inc. in the United States and other countries. Microsoft and Windows are registered trademarks of Microsoft Corporation. Red Hat is a registered trademark of Red Hat, Inc. Linux is a registered trademark of Linus Torvalds. © 2005 Dell Inc. All rights reserved.

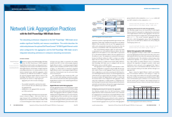
TABLE OF CONTENTS

ADVERTISER INDEX

ONLINE XTRA

DELL POWER SOLUTIONS MAGAZINE ARTICLES AVAILABLE ONLY AT WWW.DELL.COM/POWERSOLUTIONS

WWW.DELL.COM/ POWERSOLUTIONS



Network Link Aggregation Practices with the Dell PowerEdge 1855 Blade Server

By Bruce Holmes

The networking architecture integrated on the Dell PowerEdge 1855 blade server enables significant flexibility and resource consolidation. This article describes the relationship between the optional Dell PowerConnect 5316M Gigabit Ethernet switch when configured for link aggregation and the Dell PowerEdge 1855 blade server's integrated networking architecture in enterprise networking environments.

Open Source Utilities for Remote Monitoring of HPC Clusters

By Yung-Chin Fang, Randy DeRoeck, Rinku Gupta, and Monica Kashyap

Computing centers continue to scale out in a quest for additional processing power, making remote hardware management critical for cluster hardware uptime. This article provides examples of how open source utilities such as IPMItool and Ganglia can be utilized and integrated into large computing infrastructures to monitor and manage hardware health conditions.

Enhancing High-Performance Computing Clusters with Parallel File Systems

By Amina Saify; Garima Kochhar; Jenwei Hsieh, Ph.D.; and Onur Celebioglu

The performance of I/O subsystems within high-performance computing (HPC) clusters has not kept pace with processing and communications capabilities. This article discusses the advantages that parallel file systems can offer HPC clusters and how these specialized file systems can help enhance overall scalability of HPC clusters.

Designing High-Performance Computing Clusters

By Yung-Chin Fang; Saeed Iqbal, Ph.D.; and Amina Saify

HPC clusters have evolved from experimental computing architecture to mainstream supercomputing architecture. Today, HPC cluster architecture integrates multiple state-of-the-art, industry-standard components to provide aggregated, cost-effective supercomputing power. This article examines the influence of component selection on cluster performance and manageability.

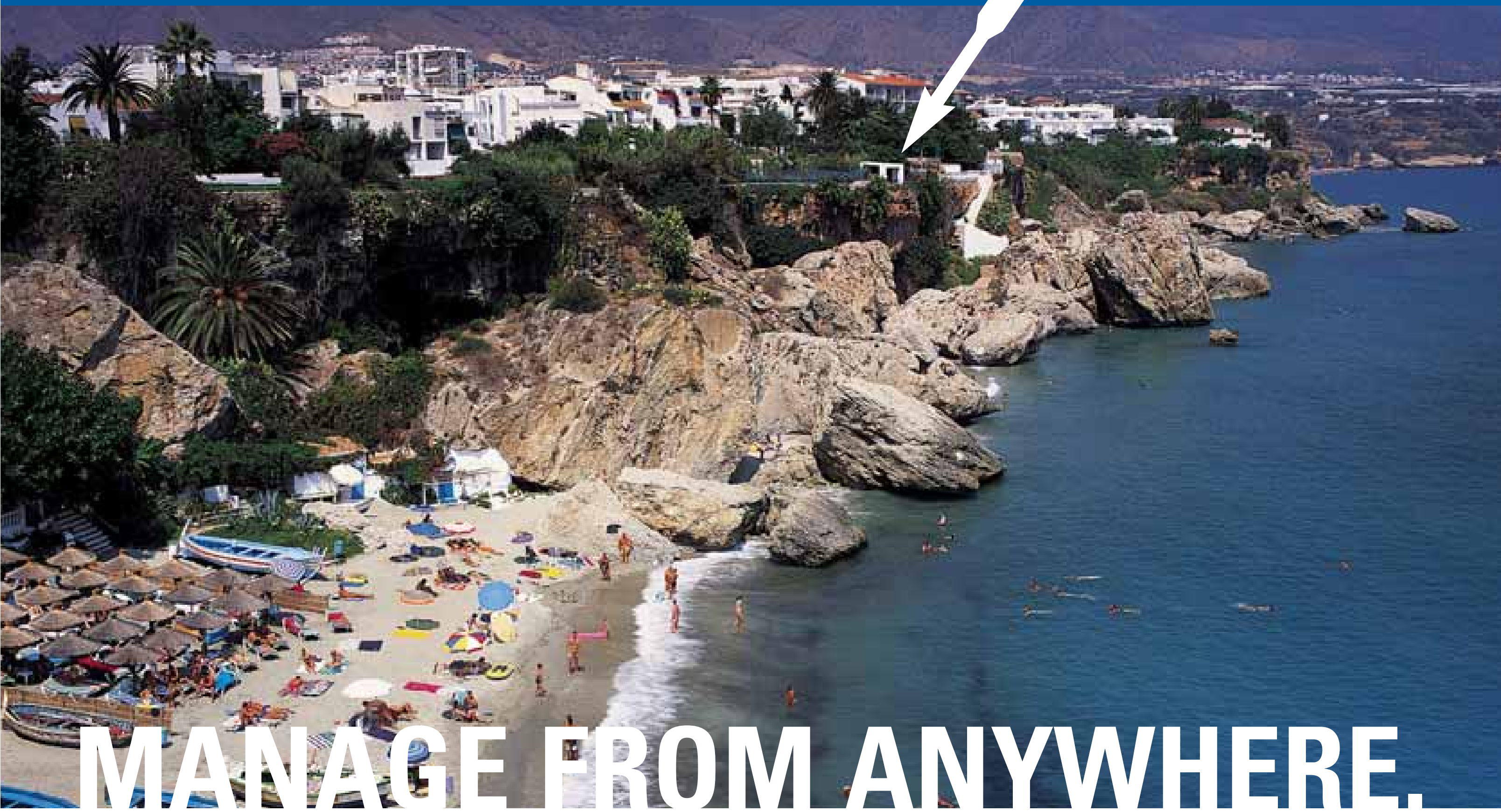
SEE IT HERE FIRST!

Check the *Dell Power Solutions* Web site for our late-breaking exclusives, how-to's, case studies, and tips you won't find anywhere else. Want to search for specific article content? Visit our Related Categories index online at www.dell.com/powersolutions.

Dave is on vacation.

He's been notified of a problem with his company's servers.
He's able to manage those servers from anywhere.
He's not cutting his vacation short.

He's right here.



The Dell remote console switch provides the traditional functionality of KVM switching with the added benefit of **KVM Over IP**. Using Dell's remote console software, simply point-and-click to take control of your servers.

Whether you're at the rack, in your office or across the globe your data center will always be within your reach.

Take **control** of your servers with
the **Dell Remote Console Switch.**

Digital Availability. **Easy as** DELL

www.dell.com

Dell and the Dell logo are trademarks of Dell Inc.

It's All About You

The last four issues of the *Dell Power Solutions* print edition have delivered 115 leading-edge technical articles spread across a voluminous 548 pages—enough to fill a sizeable book. In fact, if those 548 cumulative pages were bound into a single volume, the folio would place us squarely between the 454 pages of the best-selling *The Da Vinci Code* and the 672 pages of the forthcoming *Harry Potter and the Half-Blood Prince.* But have subscribers actually *read* this "book" of articles? That's just one of the many questions we set out to answer when we launched our first-ever reader survey.

For the initial benchmark survey in January 2005, we sent a sample of our U.S.-based subscribers an e-mail invitation to participate in the Web-based research. Enough of the e-mail messages successfully made the trek through firewalls and spam filters that a healthy 10 percent of recipients responded to the survey. Although it may not have the mass literary appeal of *The Da Vinci Code, Dell Power Solutions* was obviously a page-turner: Most survey respondents opened and actually read at least three of the last four issues. Even more telling about our readers' appetite for technical content, however, was precisely how many pages in the most recently surveyed issue respondents had read: an average 48 percent of the October 2004 issue. (If you're counting, that translates to 69 of the issue's 144 pages.)

In another section of the lengthy survey, we asked respondents to reveal the "most useful article" over the past four issues. By a wide margin, storage solution articles were ranked the most popular (26 percent of responses), followed by server solutions (14 percent), virtualization technology (8 percent), clustering (8 percent), Dell™ OpenManage™ infrastructure (6 percent), and the Linux® OS (5 percent). When asked what should be added to the magazine, respondents most often requested increased coverage on tips, how-to's, or best practices (12 percent); new product information (9 percent); and real-world case studies (8 percent)—closely followed by small to medium business articles, technology news and trends, and product comparisons. We are still wading through the megabytes of data amassed in our survey process, but your responses and free-form comments are highly valued and much appreciated. The *Dell Power Solutions* editorial team will continue to use your feedback to help shape the magazine into the most effective IT resource it can be.

Back to the present, this May 2005 issue of *Dell Power Solutions* offers a diverse set of lab-fresh technical content, starting with a strong emphasis on storage. Our cover story, "Advancing Your Enterprise Storage Strategy," focuses on new and evolving storage technology, followed by seven more articles to help round out your storage strategy—ranging from 4 Gbps Fibre Channel ("Planning for Next-Generation SANs with 4 Gbps Fibre Channel Switches") to the new Disk Data Format specification ("Solving the RAID Compatibility Puzzle"). In total, you will find 24 articles in this issue, plus four Web-exclusive articles. Also, while you are online at www.dell.com/powersolutions, check out our new Related Categories index for rapid access to our library of *Dell Power Solutions* articles. Enjoy!

Tom Kolno

Tom Kolnowski
Editor-in-Chief
tom_kolnowski@dell.com
www.dell.com/powersolutions

BACKUP
EXEC 10

Advancing Your Enterprise Storage Strategy

Dell's scalable enterprise strategy for industry-standard storage platforms is more flexible and manageable than ever. By broadening the scope of its offerings, Dell enables organizations of all sizes to extend the reach of their networked storage capabilities—providing outstanding storage platforms that are designed to deliver high availability, performance, and capacity for demanding enterprise applications.

BY VICKI VAN AUSDALL

Related Categories:

Dell PowerEdge servers

Dell PowerVault storage

Dell/EMC storage

Fibre Channel storage

File servers

iSCSI storage

Network attached storage (NAS)

Standards

Storage array

Storage

Storage architecture

Storage software

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

As more enterprises come to rely on large databases, e-commerce, and other I/O-intensive online transaction processing (OLTP) systems to conduct daily business activities, IT organizations must identify cost-effective ways to build, manage, and scale their storage infrastructures to enable high availability, efficient backup, and effective business continuity for key enterprise applications. At the same time, regulatory compliance requirements are calling upon administrators to archive ever-larger amounts of business and messaging data—and to keep that data secure and readily accessible.

Ideally, administrators would have the tools readily at hand to streamline the deployment and management of storage resources while providing the flexibility to meet service-level agreements and set priorities for archiving and accessing different types of enterprise data. However, many organizations are strapped by limited budgets, and existing IT infrastructures may fall short of the task at hand. In addition, administrators typically must work with legacy storage systems that have been deployed on an ad hoc basis by individual workgroups and branch locations. All too often, such systems are tied to a single server and application and can be either underutilized or overutilized, suffering from degraded performance caused by growing workloads.

Dell's scalable enterprise strategy offers a versatile framework that can help organizations build upon their existing IT infrastructures to align with today's business goals cost-effectively and add capacity incrementally as business needs evolve. The Dell scalable enterprise strategy is designed to simplify operations through increased standardization and automation and to help improve resource utilization through consolidation of server and storage platforms.

The building blocks of Dell's scalable enterprise strategy are modular data center components based on industry standards. Key components include server, storage, networking, and management products that are often developed in strategic alliances with industry-leading third-party hardware and software providers and then validated in various Dell and partner labs to help provide seamless interoperability and data access across a broad range of platforms.

Planning for fast, flexible business response

The Dell scalable enterprise strategy is a pragmatic, phased approach that can be geared to the needs of any size organization—including small to medium businesses (SMBs) as well as large enterprises with worldwide branch-office operations. Based on industry-standard data center components that can be upgraded incrementally, Dell™ PowerVault™ and Dell/EMC networked storage products range from low-cost network attached storage (NAS) and Internet SCSI (iSCSI) arrays to high-performance Fibre Channel storage area networks (SANs). These building blocks can help administrators configure flexible, cost-effective enterprise storage networks that support diverse, far-reaching capacity and performance requirements. (For information about Dell's latest server systems, see "Driving databases with the new Dell PowerEdge 6800 and PowerEdge 6850 servers" in this article.)

Extending the reach of networked storage for SMBs

To help budget-conscious organizations extend the reach of their networked storage capabilities, Dell and EMC have partnered to introduce the Dell/EMC AX100i array, an iSCSI-based storage array that is designed to provide up to 3 TB of storage capacity (see Figure 1). Because the iSCSI standard allows data to be transmitted over existing Gigabit Ethernet networks as well as IP-based networks, the AX100i array gives administrators the flexibility to configure networked storage resources without investing in a separate Fibre Channel storage fabric.

As a result, the AX100i array enables SMBs to take their first steps toward consolidating storage platforms and building scalable enterprise storage infrastructures. Storage consolidation also helps administrators to streamline operations and improve resource utilization, particularly compared to direct attach storage (DAS) configurations that may have been implemented on an ad hoc basis. In addition, the AX100i array can support fast, flexible expansion of storage capacity by connecting to as many as eight storage servers over a standard IP-based LAN, metropolitan area network (MAN), or wide area network (WAN).

Figure 1. Dell/EMC AX100i storage array

Integrating iSCSI and NAS platforms with Fibre Channel SANs

The use of iSCSI storage is by no means limited to SMBs. IP-based and Fibre Channel–based storage platforms can coexist and provide complementary functions in any size organization to help advance overall goals of storage consolidation, enhanced utilization, and simplified operations. In fact, NAS—a traditional approach to cost-effective storage that predates iSCSI—can also be connected to SAN environments.

Typically, organizations implement two types of storage based on the differing needs of their application environments: NAS to support applications that lend themselves to file-based I/O, and SAN or iSCSI to support applications that lend themselves to block-based I/O. However, configuring separate storage architectures in the IT infrastructure can create management complexity, require additional administrative resources, and lead to underutilization of IT investments.

IP-based and Fibre Channel–based storage platforms can coexist and provide complementary functions in any size organization to help advance overall goals of storage consolidation, enhanced utilization, and simplified operations.

The recently announced Dell/EMC NS500G system is a NAS gateway that can help keep total cost of ownership low by allowing file-level information to be stored on a Fibre Channel SAN. In this way, the Dell/EMC NS500G gateway can help optimize SAN investments and enhance the efficiency of storage management by consolidating all file and block data in one storage pool. A NAS gateway also enables administrators to cost-effectively use existing NAS platforms, such as the Dell PowerVault 745N storage server, for enterprise infrastructure purposes such as disk-to-disk backups that help shrink backup windows and expedite recovery tasks.

Enhancing storage infrastructure with Fibre Channel SANs

Whether the business driver is data growth or fast, transaction-based enterprise systems, a Fibre Channel SAN can help administrators respond quickly to enterprise requirements for enhanced I/O performance, storage capacity, and application availability. At the heart of the SAN is the storage array, and Dell/EMC storage arrays are designed to provide versatile data center building blocks that can be scaled incrementally whenever and wherever they are needed.

The Dell/EMC family of Fibre Channel storage arrays includes the 2 Gbps Dell/EMC AX100, CX300, CX500, and CX700 storage platforms. The entry-level AX100 array can provide cost-effective

Fibre Channel SAN performance and availability while offering a strong migration path to higher-capacity Fibre Channel arrays in the CX series. Like the AX100i, its iSCSI counterpart, the AX100 is designed to provide up to 3 TB of storage capacity.

Each Dell/EMC CX series array is designed to provide data protection through redundant hardware. The CX300 is designed to support day-to-day production activities and to anchor medium-size SAN deployments, and it can provide up to 13.4 TB of data storage. For large SANs with workloads such as high-performance Web-serving and data-warehousing applications, the midrange CX500 can provide up to 27 TB of storage. The CX700 platform offers the most powerful processing resources available in the Dell/EMC CX series of Fibre Channel arrays and provides up to 58.4 TB of storage capacity—making the CX700 (see Figure 2) suitable for compute-intensive databases and OLTP applications.

Deploying powerful, scalable, industry-standard data center components such as the CX500 and CX700 can help administrators eliminate isolated islands of storage that may be scattered throughout the organization and help improve resource utilization by consolidating storage capacity onto fewer higher-performance storage modules. The CX300, CX500, and CX700 arrays are also designed to interoperate with legacy Dell/EMC CX systems to provide investment protection.

Looking ahead to 4 Gbps SAN technology

Emerging 4 Gbps SAN products are being designed to enhance network throughput for demanding enterprise workloads. The 4 Gbps devices are expected to include auto-sensing capabilities that will enable them to recognize 1 Gbps and 2 Gbps devices on the Fibre Channel network and operate at the appropriate speed. Such plans for backward compatibility can allow administrators to advance the throughput capabilities of the enterprise storage infrastructure while keeping 1 Gbps and 2 Gbps devices in service to protect existing IT investments.

Several major hardware vendors have already announced 4 Gbps SAN products, including Brocade. The Brocade SilkWorm 4100 switch—the industry's first 4 Gbps Fibre Channel switch—is designed to combine 4 Gbps technology with high-availability features that support mission-critical SAN environments.[1] Based on the same standards as 1 Gbps and 2 Gbps Fibre Channel technology, the SilkWorm 4100 switch is designed to be compatible with current 2 Gbps Dell/EMC arrays, allowing administrators to incorporate the 4 Gbps Brocade hardware into the storage infrastructure to facilitate higher throughput when Dell and EMC begin releasing 4 Gbps arrays.

Figure 2. Dell/EMC CX700 storage array

Simplifying operations and optimizing resource utilization

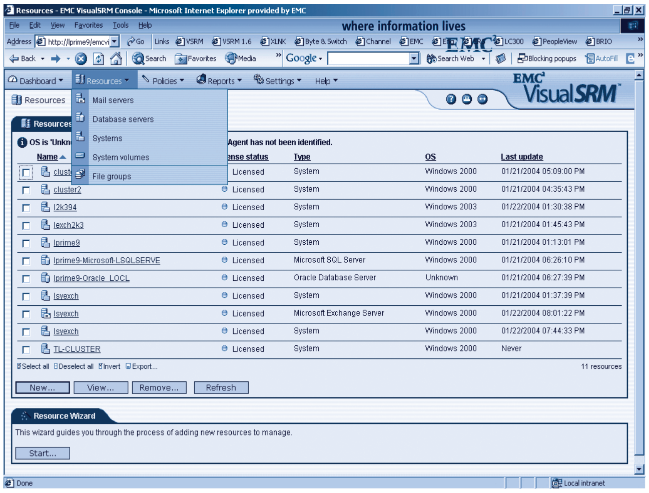
To use the enterprise storage infrastructure to best advantage, administrators must have the capability to balance workloads effectively across many and varied storage resources and to manage those resources efficiently. As enterprises grow, the size and complexity of storage infrastructures can turn the simplest management tasks into an overwhelming burden for administrators. Using SANs to enable the consolidation of storage platforms across the enterprise can help organizations optimize both administrative and storage resources.

To truly benefit from storage consolidation, however, administrators need a common set of tools to view and manage centralized storage. For example, the EMC® Navisphere® Management Suite allows administrators to provision storage capacity across the enterprise SAN. Navisphere is designed to provide flexible provisioning and central control over Dell/EMC CX300, CX500, and CX700 arrays as well as previous-generation Dell/EMC CX series arrays. Meanwhile, Navisphere Express can be used to manage Dell/EMC AX100 and AX100i arrays, and includes a quick setup wizard especially useful for SMBs.

From the Navisphere console, administrators can provision additional storage capacity without taking the system offline—to help ensure that unpredicted spikes in capacity demand can be met immediately, without disruption to the production environment. Navisphere can also simplify management tasks by helping to automate the process of configuring storage systems. The Navisphere suite enables administrators to save time and prevent errors by allowing the use of templates to duplicate similar environments. The suite can also be set up to send automatic alerts when a change in system status occurs, helping administrators identify problem conditions before failures occur and proactively direct storage resources where they are needed to keep systems up and running.

Complementing Navisphere functionality, EMC VisualSRM™ software can help enterprises move from reactive storage resource allocation toward automated management even in multivendor, multiplatform environments. VisualSRM is designed to provide

[1] For more information about the Brocade SilkWorm 4100 switch, see "Planning for Next-Generation SANs with 4 Gbps Fibre Channel Switches from Brocade" by Spencer Sells in *Dell Power Solutions,* May 2005.


Resources - EMC VisualSRM Console - Microsoft Internet Explorer provided by EMC
File Edit View Favorites Tools Help
where information lives
Address http://lprime9/emcvi
Go
Links VSRM VSRM 1.6 XLNK Byte & Switch Channel EMC LC300 PeopleView BRIO
Back Search Favorites Media
Google
Search Web
Blocking popups
AutoFill
Dashboard
Resources
Policies
Reports
Settings
Help
EMC² VisualSRM
Resources
Mail servers
Database servers
Systems
System volumes
File groups
Select all Deselect all Invert Export...
New... View... Remove... Refresh
Start...
Done
Local intranet

Resources

OS is 'Unkn… Agent has not been identified.

	Name	…ense status	Type	OS	Last update
□	clust…	Licensed	System	Windows 2000	01/21/2004 05:09:00 PM
□	cluster2	Licensed	System	Windows 2000	01/21/2004 04:35:43 PM
□	l2k394	Licensed	System	Windows 2003	01/22/2004 01:30:38 PM
□	lexch2k3	Licensed	System	Windows 2003	01/21/2004 01:45:43 PM
□	lprime9	Licensed	System	Windows 2000	01/21/2004 01:13:01 PM
□	lprime9-Microsoft-LSQLSERVE	Licensed	Microsoft SQL Server	Windows 2000	01/21/2004 06:26:10 PM
□	lprime9-Oracle_LOCL	Licensed	Oracle Database Server	Unknown	01/21/2004 06:27:39 PM
□	lsyexch	Licensed	System	Windows 2000	01/21/2004 01:37:39 PM
□	lsyexch	Licensed	Microsoft Exchange Server	Windows 2000	01/22/2004 08:01:22 PM
□	lsyexch	Licensed	System	Windows 2000	01/22/2004 07:44:33 PM
□	TL-CLUSTER	Licensed	System	Windows 2000	Never

11 resources

Resource Wizard

This wizard guides you through the process of adding new resources to manage.

Figure 3. Resource monitoring and management using EMC VisualSRM

a consolidated view of the capacity utilization of hosts, file systems, and storage devices on the SAN and an integrated view of how resources are being used (see Figure 3). These capabilities enable administrators to automate disk space management and capacity planning, discover and classify data across the enterprise, and then take immediate action—for example, compressing outdated files or deleting temporary files. By allowing administrators to identify storage usage patterns and establish automated storage policy goals accordingly, VisualSRM can help organizations reclaim wasted disk space and optimize storage planning.

Developing effective data backup and compliance strategies

Planning is essential for the proper management of backups and archiving to help avoid affecting the performance of business-critical applications on the enterprise SAN. Dell's scalable, standards-based disk and tape storage components can be used to address backup, archiving, and regulatory compliance issues.

Accelerating and automating the backup process

The requirement to back up and retain ever-increasing business and messaging data can significantly lengthen disk-to-tape backup processes. For enterprises whose daily backup processes threaten to spill over into production time, cost-effective disk-based Dell/EMC storage arrays can help enable disk-to-disk storage. Because disks are generally faster than tape, the use of disk-to-disk backups can accelerate backup and recovery times.

Tape continues to be a strategic component of enterprise data protection strategies, and organizations can benefit from the portability and security of archived tape for long-term data retention as well as disaster recovery. Tape autoloaders such as the Dell PowerVault 122T and tape libraries such as the PowerVault 132T and PowerVault 136T can automate manual tape-handling processes to free administrators for other tasks and help eliminate human error in tape backup processes. Dell PowerVault tape systems combined with data archiving and backup software from leading software providers—including VERITAS Backup Exec and NetBackup, Yosemite TapeWare, LEGATO NetWorker, and CommVault Galaxy—can enable organizations of any size to choose the appropriate backup and data protection strategy.

Enabling secure long-term storage and regulatory compliance

To store increasing amounts of data cost-effectively over the long term, administrators must cultivate the organizational awareness that not all data is of equal value or needs to be accessed with equal frequency. In fact, storing all data the same way can be both risky and expensive: Organizations may be wasting high-priority disk space on low-value information or stashing critical data where it cannot be easily located, accessed, and protected. By matching data value to storage costs and levels of protection, enterprises of all sizes can help speed access to their most critical information, minimize the risk of data loss, and build a cost-efficient storage infrastructure. For short-term archiving or storage of data that must be frequently accessed, secondary disk storage is often an effective choice because disk access is random by nature and generally quicker than sequential tape access. Tape affords a cost-effective alternative for data that is largely inactive.

Recently, methods have emerged to help organizations meet stringent compliance regulations for storing and accessing *fixed content* such as check images, patient records, and e-mail messages—digital assets that do not change over time yet must be kept secure and accessible for lengthy periods to facilitate audits and response to litigation. Write once, read many (WORM) devices such as the EMC Centera™ system are designed to streamline the storage and retrieval of fixed content and to scale cost-effectively. In addition, data archiving software tools such as VERITAS Enterprise Vault and Discovery Accelerator help enable quick search and retrieval of archived data, allowing enterprises to respond quickly to requests for archived documents and messages.

To store increasing amounts of data cost-effectively over the long term, administrators must cultivate the organizational awareness that not all data is of equal value or needs to be accessed with equal frequency.

DRIVING DATABASES WITH THE NEW DELL POWEREDGE 6800 AND POWEREDGE 6850 SERVERS

The four-processor Dell PowerEdge™ 6800 tower and PowerEdge 6850 rack servers are designed to provide enterprise-class computing at an entry-level price. The new systems are geared to address the performance challenges of database, e-commerce, OLTP, and other enterprise applications—such as virtualization—that can benefit from powerful multiprocessing, scalable memory, and large I/O bandwidth. As organizations begin migrating to demanding 64-bit applications, the PowerEdge 6800 and PowerEdge 6850 servers can enable organizations to run resource-intensive applications on industry-standard components and software—thereby helping to avoid the cost and potential management complexity of proprietary technology.

To provide high performance, scalability, and availability, both the PowerEdge 6800 and PowerEdge 6850 servers incorporate up to four 64-bit Intel® Xeon™ processors MP, which support Intel Extended Memory 64 Technology (EM64T); up to 8 MB of level 3 (L3) cache; a 667 MHz frontside bus (FSB); double data rate 2 (DDR2) memory; and hot-pluggable hard drives, cooling fans, and power supplies. The servers have been tested and certified with leading operating systems and software applications including Microsoft® Windows Server™ 2003, Enterprise Edition; Red Hat® Enterprise Linux®; Microsoft SQL Server 2000 Enterprise Edition; and Oracle® Database 10*g*.

For enhanced manageability, the PowerEdge 6800 and PowerEdge 6850 servers take advantage of Dell management software and hardware to streamline administrative tasks. For example, the Dell OpenManage™ 4.3 infrastructure helps automate discovery and updates of system software, while the Dell Remote Access Card 4 (DRAC 4) helps simplify operations and extend administrative reach by allowing IT staff to manage servers independently of the server's OS state, without having to physically access the hardware. All together, these features can enhance IT efficiency and help improve business response.

Additionally, for easier integration with Fibre Channel storage networks, the rack-mountable PowerEdge 6850 server (see Figure A) can be equipped with an optional integrated, dual-port host bus adapter (HBA). This QLogic-based daughter card, the Dell 2362M, does not require a slot. Standard Fibre Channel HBAs can also be installed in the Peripheral Component Interconnect Extended (PCI-X) slot and in high-performance PCI Express slots.

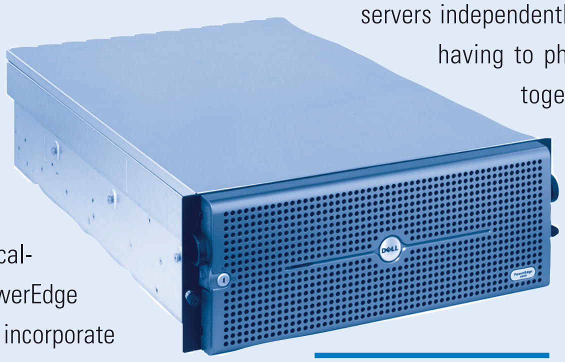


Figure A. Dell PowerEdge 6850 rack server

Fostering standards and interoperability to advance the scalable enterprise strategy

Dell supports standards that help storage hardware and management software from different vendors to interoperate smoothly. For example, Dell is actively involved with the Storage Networking Industry Association (SNIA)[2] efforts to develop specifications that are designed to allow data to be stored and accessed seamlessly across heterogeneous storage platforms.

By taking a leadership role in standards initiatives and working with partners in storage and other areas to develop scalable, industry-standard data center components, Dell is helping to advance the goals of the scalable enterprise. These initiatives are designed to enable organizations of any size and scope to create a flexible IT infrastructure that allows quick response to immediate business needs and supports future business development by providing highly scalable computing, storage, and networking capabilities.

Vicki Van Ausdall is a senior editor at *Dell Power Solutions,* with 12 years of experience as a technical writer and editor for various high-tech publications in the San Francisco Bay Area. She has a B.A. in English Literature from Hamilton College.

[2] For more information about Dell's support for storage industry standards such as the Common RAID Disk Data Format (DDF) specification, see "Solving the RAID Compatibility Puzzle" by Matthew Brisse and Bill Dawkins, Ph.D., in *Dell Power Solutions,* May 2005.

Emulex.
We network storage

Solving the RAID Compatibility Puzzle

The Common RAID Disk Data Format (DDF) specification, recently published by the Storage Networking Industry Association, is designed to alleviate compatibility issues in RAID installations. This article discusses the importance and implications of the DDF storage specification.

BY MATTHEW BRISSE AND BILL DAWKINS, PH.D.

Related Categories:

Common RAID Disk Data Format (DDF)

File systems

Industry standards

RAID

Storage

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

For many years, RAID has been a critical component of server and storage systems. However, even when two array vendors implement RAID-5 in a similar fashion—such as a disk set with distributed parity—one array will not necessarily be able to read data from a RAID-5 set created by the other vendor's array. This incompatibility is a result of how configuration information is stored on each disk within a RAID subsystem. The configuration information is in the form of a data structure and is called a configuration on disk (COD). A COD typically includes assignment of physical disks to a RAID group, RAID levels of RAID groups, data mapping on RAID groups, and hot-spare assignments. Currently, each vendor implements its own proprietary COD format, and in some cases, a single vendor will implement different CODs for the different RAID products within its own portfolio. Different CODs prevent a vendor from determining the RAID format used on a RAID set created by another vendor.

The typical data center has many different RAID implementations, from internal software-based RAID approaches that are included with an OS to off-the-shelf RAID products. Because COD structures differ, disks cannot be physically moved from one RAID system to another with the data in place. In addition, it is difficult to reuse RAID disks from retired or obsolete hardware because RAID disks cannot be moved between vendor implementations with the data in place. As a result, it can be expensive to consolidate or redistribute data among RAID systems in the data center as performance or capacity demands dictate—and to scale small implementations to larger, flexible configurations. This situation can negatively affect return on investment.

In many cases, data loss or corruption can occur during the migration process. For example, if an administrator moves disks from one vendor's RAID system to a system from a different vendor, the disk format most likely will be unrecognized by the new controller and possibly perceived as blank. In addition, disk signatures may be written to disk, destroying the data. Administrators may inadvertently create a new RAID group over the existing data, resulting in loss of data. The lack of RAID interoperability is an industry-wide challenge for administrators and vendors alike.

RAID challenges that DDF addresses

To address this concern and help facilitate flexibility and choice for system administrators, the Storage Networking Industry Association (SNIA) Common RAID Disk Data Format (DDF) Technical Working Group (TWG) has developed the Common RAID DDF specification. The DDF specification standardizes the RAID configuration data structure and describes how data is formatted across disks in a RAID group. This specification enables a basic level of interoperability among RAID systems from different vendors.

The DDF data structure describes how data is distributed across the disks in a RAID group. Figure 1 shows the scope of the DDF data structure in the context of the SNIA Shared Storage Model. The DDF scope is limited to the interface between the block aggregation implementation and storage devices, as defined in the SNIA Shared Storage Model. DDF does not standardize the RAID controller interface of the OS or create a single driver. Also, the DDF structure addresses server-resident RAID implementations—such as Peripheral Component Interconnect (PCI) RAID cards and RAID on Motherboard (ROMB). The DDF TWG is currently focused on direct attach storage (DAS), and the scope of the specification does not include high-end external RAID configurations.

Part of the DDF specification defines how data is distributed for several basic RAID levels. The specification uses mathematical formulas to document how data is formatted for RAID levels defined by the DDF structure to help prevent misinterpretation. The DDF data structure also allows RAID groups with vendor-unique RAID data formats. When a RAID group containing a unique format is moved to a different RAID system that does not support the format, the DDF specification is designed to enable the different system to read the DDF structure stored on each disk and detect the noncompatible RAID data formats. Potential data loss can be prevented because the different RAID system will not overwrite the data without administrator confirmation.

The DDF structure resides on every disk behind a RAID controller, providing redundancy and helping to prevent random disks from being mixed into a RAID group from a different configuration. Collectively, the DDF structure on the disks defines how data is distributed in a RAID group. The DDF structure describes features such as the controller data, physical disk data (such as online state), virtual disk data (such as initialization state), hot-spare assignments, bad-block management, and vendor-specific logs. The DDF structure also describes virtual disk configurations such as RAID levels, participating disks, stripe sizes, and cache policies.

Where DDF can enhance interoperability

The Common RAID DDF specification is designed to help address far-reaching compatibility issues affecting internal ROMB, disaster recovery, backup to disk, and technology transition.

Internal ROMB. Servers shipped with RAID implemented on the motherboard may allow RAID data formats to be applied to the server's internal disks. As the server's data set grows, administrators often find they need to move the internal disks and their data to a larger DAS configuration with external JBOD (Just a Bunch of Disks) storage. One method of migration is to back up the RAID group, transfer the disks to the new storage system, reconfigure the disks behind the new RAID controller, and restore the data. However, this procedure is time-consuming and can increase the risk of data loss.

Disaster recovery. Many administrators have spare servers that can be used if a primary server crashes. With a server-based ROMB approach, administrators must maintain the same or compatible RAID implementations in their server inventory. Otherwise, restoring data on a replacement server from a tape backup might take hours. This problem is compounded when a data center contains servers from multiple vendors. Ideally, administrators would be able to exchange any server and know that the RAID implementation from vendor *X* would work with the RAID implementation from vendor *Y*. Organizations could then economize by retaining fewer servers in their data centers.

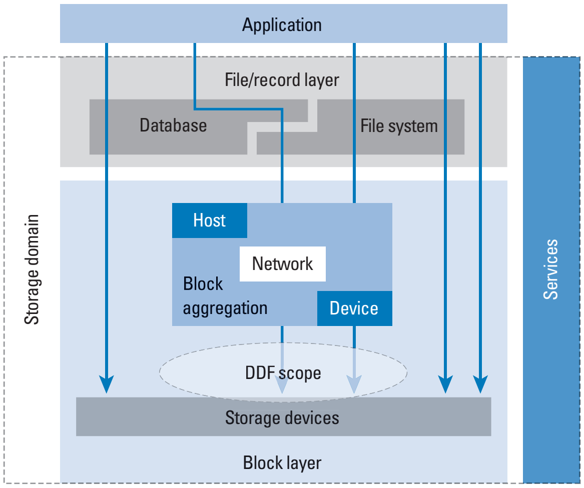


Figure 1. DDF scope in the SNIA Shared Storage Model

Backup to disk. If administrators perform a backup to disk, take the disks offline for archiving, and then try to read the disks from a different server, the RAID implementations must be 100 percent compatible. Ensuring compatibility between RAID configurations over time can be a daunting task.

Technology transition. When an administrator replaces an older server with storage comprising an internal PCI Extended (PCI-X) RAID controller and external JBODs, the administrator may need to migrate the data to a newer system with PCI Express slots. To take advantage of the performance benefits of PCI Express, the administrator would need to install a new RAID controller. Without a common RAID data format, additional steps would have to be taken to migrate the data.

The future of DDF

The SNIA's goal is to be a catalyst for standards that benefit enterprises, and the DDF specification is the SNIA's first approved host-based storage architecture. Currently, the DDF project has been submitted to the InterNational Committee for Information Technology Standards (INCITS) to begin the ANSI standardization process. DDF-compliant RAID products are expected to be available in 2006.

Matthew Brisse is a member of the Chief Technology Office for Dell and is the vice chair of the SNIA Board of Directors.

Bill Dawkins, Ph.D., is a member of the Chief Technology Office for Dell and is the chair of the SNIA's DDF Technical Working Group.

FOR MORE INFORMATION

Common RAID DDF:
www.snia.org/tech_activities/ddftwg

Advanced Storage Life-Cycle Techniques for

Dell/EMC SAN MetaLUN Expansion

The introduction of metaLUNs in EMC® FLARE™ Release 12 software allows organizations to extend capacity, reliability, and performance beyond the realm of a standard RAID group. Using metaLUNs, administrators have the opportunity to account for multiple variables in a way that can help shape the storage design into a robust, modular, high-performance system. This article discusses best practices for using a metaLUN—starting from its initial creation in the RAID group layout—to enable a virtual storage architecture to be designed effectively for the duration of both short-term and long-term storage life cycles.

BY ARRIAN MEHIS

Related Categories:

Dell/EMC storage

Fibre Channel

metaLUNs

Performance

RAID

Storage

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

Enterprise Dell/EMC storage area networks (SANs) are often characterized by high, multi-terabyte capacity; a large host count; and consequently high expectations and demands for performance. Storage architects and administrators will find that strategically designed metaLUNs can provide both short-term and long-term storage solutions. MetaLUNs can be used for all types of workloads—including transactional databases, file sharing, media streaming, and backup servers—and for any logical disk that is mapped to an EMC FLARE logical storage unit (LUN). This article provides a step-by-step strategy to help achieve a robust storage back end for both short-term and long-term storage life cycles.

Defining the objectives of metaLUN use

The first step is to determine what administrators want to achieve by implementing metaLUNs—this step is the building block of an effective metaLUN architecture. The common goals are efficient use of capacity, flexible expansion of devices, high performance, and consistent performance as capacity grows. Typical data access types can be divided into two distinct categories: random and sequential. Workloads may be mixtures of both types, containing varying percentages of reads and writes.

Random access consists of unpredictable bursts of data access patterns. For the most part, the data prefetch and read cache mechanisms have little benefit, if any, for

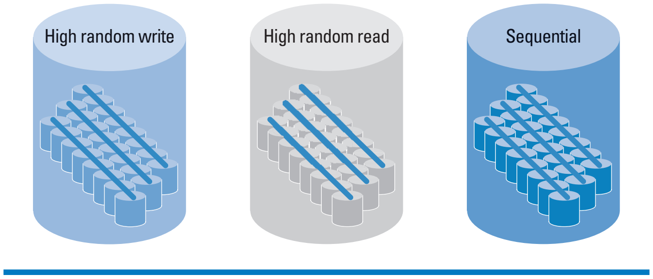


Figure 1. Organizing disks by access type

random access. Administrators should design storage architectures for the most effective distribution of these localized random bursts over as many spindles as possible; there should never be an idle disk. When the load of any application is distributed over all available drives, the high spindle count acts as a cushion while latency is decreased substantially. That is, the large pool of drives allows the storage processor cache to de-stage quickly. This approach allows the data residing in the cache to flush efficiently to the disk drives.

Sequential access consists of predictable data access patterns across a small region of the platter. Data prefetch and read cache mechanisms are a tremendous benefit to sequential access because of the high probability that, if an element was just referenced, neighboring elements will be loaded into the cache before the next read request. The read can then take place from the cache instead of from the disk, reducing read latency. A medium spindle count is important for sequential access. Unlike random access workloads, high spindle counts are not critical for sequential access workloads. However, the disks intended for sequential access should be fenced off from the disks that are intended for random access. RAID groups that share both random and sequential access patterns will be less effective at optimizing either one.

The limitations of metaLUNs depend on the model. With the Dell/EMC CX700 storage array, up to 32 LUNs can reside in a striped component, and up to 16 components can reside in one metaLUN. Best practices dictate choosing more spindles over rack density. For example, suppose the target capacity for a logical disk is roughly 300 MB. Depending on the intensity of the workload and the data access pattern, the LUN should be created from eight 36 GB spindles as opposed to four 73 GB spindles. Again, distributing the workload over multiple spindles helps to absorb any sudden burst in activity and to maintain consistent performance.

Understanding high-level metaLUN design

After identifying data access types, administrators can organize disks into several large groups, one for each access pattern (see Figure 1). For example, three disk categories can be created, each with the appropriate RAID type:

- **High random write:** Example applications include online transaction processing and e-mail.
- **High random read:** Example applications include file systems and e-mail.
- **Sequential:** Example applications include decision support systems, media servers, and backup systems.

Administrators can create these categories simply by using RAID groups. However, RAID groups do not allow I/O profiles to be distributed over all the available drives, whereas metaLUNs do.

One key factor in metaLUN design is stripe size. A LUN stripe size is the number of effective drives times the stripe element size. So, if a component LUN has the default stripe element size of 128 blocks (64 KB), then a 2 + 2 RAID-10 LUN would have a stripe size of 128 KB (two effective drives times 64 KB).

The data for a metaLUN is striped across the LUNs at a set depth. The depth of the metaLUN stripe is the base LUN's stripe size times the metaLUN element size multiplier. So, a metaLUN having an arbitrary number of 2 + 2 RAID-10 components (each has a stripe size of 128 KB) would have a stripe depth of 512 KB (metaLUN size multiplier, which is 4, times 128 KB).[1]

For consistent performance, the stripe element size and the number of disks should be the same for each component LUN within the metaLUN (see Figure 2). This allows for consistent access

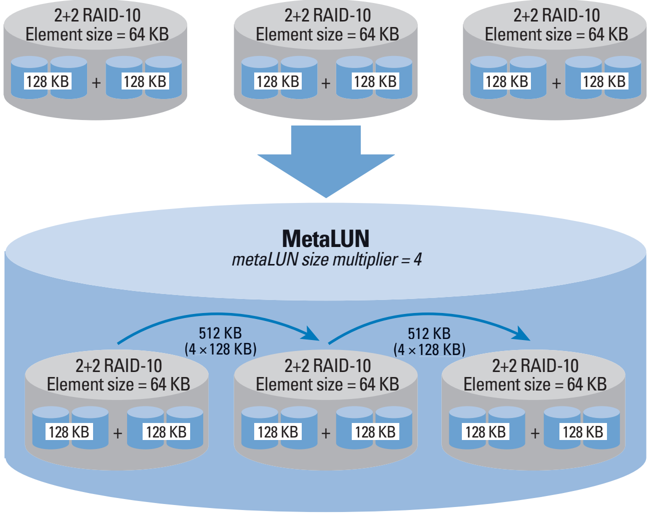


Figure 2. Using a consistent FLARE LUN stripe size for optimal traversing across the metaLUN stripe

[1] The EMC Navisphere® default metaLUN element size multiplier of 40 is too deep, and EMC plans to change this default value in a future FLARE software release. For best practices, use a multiplier of 4; it can be tuned later for unique high-bandwidth workload requirements.

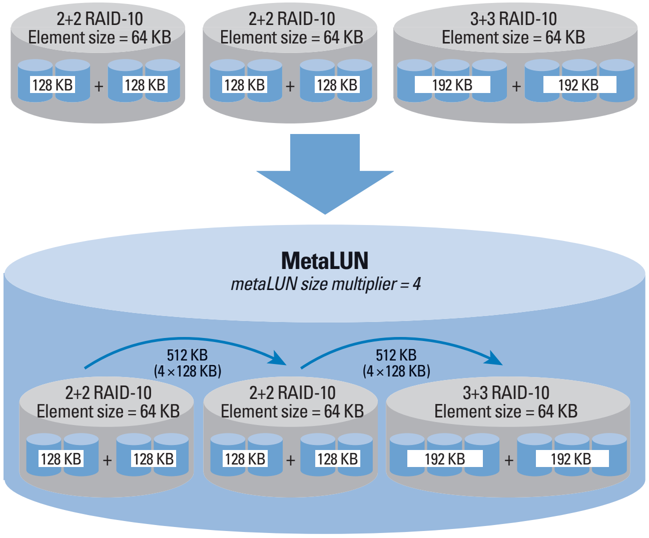


Figure 3. Using different FLARE LUN stripe sizes—metaLUN stripe depth is the same despite different individual stripe depths

across all disks because the metaLUN stripe is traversed by different I/O sizes, numbers of I/Os, and access types. For any burst of I/O within a small locality, the same number of drives will be servicing the burst. Because the metaLUN stripe depth is calculated from the first base LUN, all LUNs added to the metaLUN will have the same metaLUN stripe depth, regardless of individual LUN stripe depth (see Figure 3). However, for maximum performance and expansion capacity, each component should be constructed like the base component.

Physical metaLUN design

Consider an example scenario involving a fully populated Dell/EMC CX500 storage array with 120 drives, with half of the available drives deployed and the remaining half reserved for expansion. Two I/O access types are required for this scenario: large, sequential read and small, mixed write. By fencing the access patterns, administrators can use RAID-5 for the large, sequential read workload because it handles high bandwidth well; RAID-10 can be utilized for the small, mixed write workload because the I/O size is small (capacity is not an issue) and RAID-10 is effective at handling random I/O. To create small, manageable RAID groups, administrators can configure six groups of five-disk RAID-5 groups and seven groups of four-disk RAID-10 groups, leaving two spindles available as hot spares (see Figure 4).

By using metaLUNs as opposed to just RAID groups, administrators enable read I/O bursts for the RAID-5 groups to be absorbed by 30 spindles instead of, at most, 16 spindles. Similarly, the write I/O bursts for the RAID-10 groups will be absorbed by 28 spindles.

Expanding RAID groups

Expansion is the final test of a well-designed storage system. Ultimately, administrators will design for expansion that delivers instant capacity with little I/O impact on the system while maintaining consistent performance. Of the two methods that allow for expansion of a metaLUN, striping is the most performance effective but can be time-consuming to complete—re-striping of metaLUN data takes approximately 100 GB/hour. In addition, the metaLUN expansion incurs a dual impact on I/O performance: the RAID group expansion itself and the re-striping of data on the metaLUN. Moreover, the expansion LUN sizes must match the size of the base LUN on the metaLUN. The alternative expansion method is concatenation, which delivers instant capacity but does not distribute bursts as effectively as striping.

Storage architects and administrators will find that strategically designed metaLUNs can provide both short-term and long-term storage solutions.

However, the best practice is to combine these expansion methods by concatenating a striped component. By increasing the number of drives—with either additional drives inserted into the system or unallocated existing drives—administrators can expand the current RAID groups over the additional drives. Following the preceding recommendations of maintaining the same RAID group type, spindle count per RAID group, and stripe size, administrators can bind a new LUN in each group over the additional drives. Finally, they can concatenate the new striped LUNs to the metaLUN.

This method effectively keeps the existing data intact with no need for re-striping. Administrators increase capacity by adding a striped component to an existing metaLUN. The same disk distribution as the original component is maintained, while the online performance impact of expansion is minimized. Figure 5 shows how expansion LUNs can be added to existing RAID groups.

MetaLUNs use logical block addresses (LBAs) to access each sequential LUN in the metaLUN group. As in the scenario

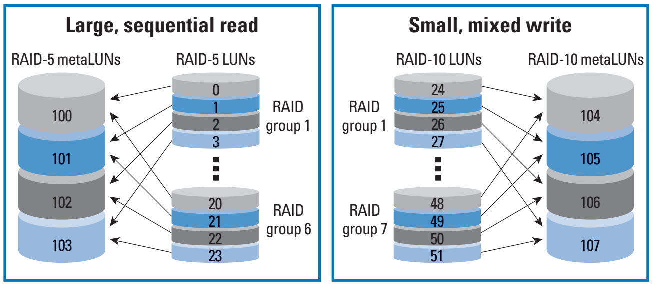


Figure 4. Creating metaLUNs that are distributed across as many RAID groups as possible set aside for particular access patterns

The first step is to determine what administrators want to achieve by implementing metaLUNs——this step is the building block of an effective metaLUN architecture.

previously described, for optimal performance, the striped LUNs are configured in different RAID groups; administrators should not stripe LUNs from the same RAID groups in a metaLUN. Because the metaLUN has a consistent 512 KB stripe, the LBA "jump" is 1,024 blocks to the next LUN. When the end of the existing metaLUN is reached, the concatenation to the striped expansion LUN is just another 1,024-block jump. Figure 6 shows how a consistent stripe size across the metaLUN is implemented as well as how a striped expansion LUN is concatenated to an existing RAID group.

Using best practices for storage architecture design

Each user workload and expectation level is unique. However, when designing an optimal storage architecture, administrators can follow the basic guidelines presented in this article:

- For best performance and reliability in expansion, use RAID-5 and RAID-10. Avoid using RAID-1 for expansion. RAID-0 can be used for temporary space when data loss is acceptable.
- Use small RAID groups for modularity, expansion, and quick rebuild turnaround. Doing so helps minimize the risk that a single disk failure will bring down the entire array because the failure is isolated to a small RAID group that can be rebuilt quickly.
- Categorize each RAID group based on I/O profile.

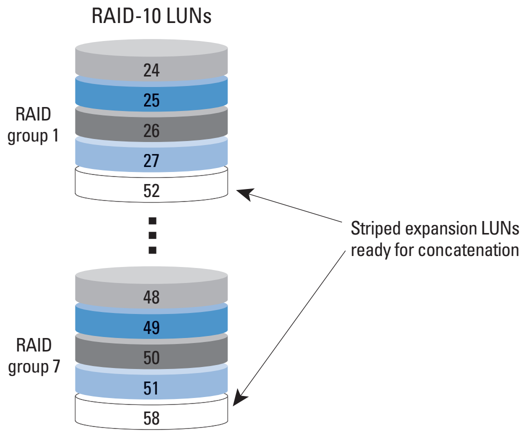


Figure 5. Expanding existing RAID groups and creating striped LUNs for concatenation to corresponding metaLUNs

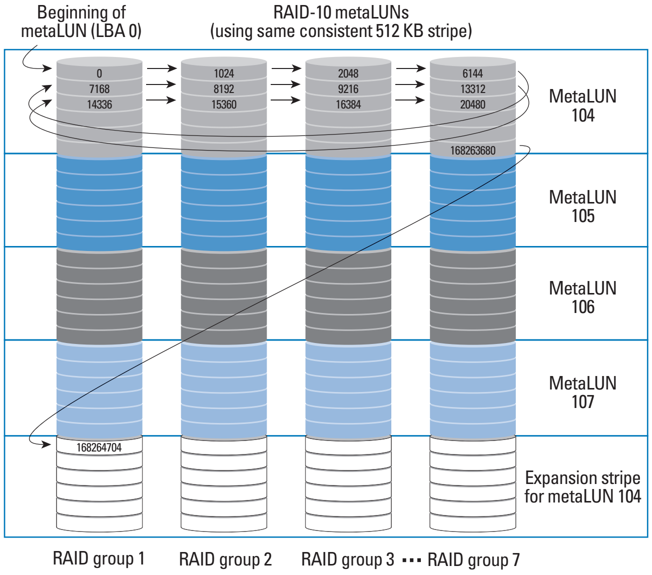


Figure 6. The mechanics of metaLUN striping and how concatenation of an expansion stripe works

- Balance metaLUNs across storage processors (SPs). The SP that owns the original base LUN used to create the metaLUN will automatically assume ownership of each component LUN added to the metaLUN.
- Use a metaLUN size multiplier of 4. Adjust for unique I/O patterns if needed.
- The metaLUN "signature" is written on the base LUN used to create the metaLUN. Avoid using LUNs from the same RAID group for every metaLUN created. Instead, rotate the base LUNs across different RAID groups for each metaLUN.
- To expand a RAID group, use the technique of concatenating a striped LUN.

These guidelines and the examples discussed in this article can help prepare administrators to build a storage system that is designed to provide high levels of performance, capacity expansion, and availability.

References

Zeryck, Dave. "Effective Use of CLARiiON MetaLUNs for the Storage Lifecycle." EMC Corporation. March 16, 2004.

Arrian Mehis is a systems engineer on the Server and Storage Performance team in the Dell Enterprise Product Group. His current focus is single-node and high-availability cluster performance analysis of Microsoft® Exchange Server on Dell™ servers and SANs. Arrian has a B.S. in Computer Engineering with a minor in Information Systems from the Georgia Institute of Technology.

Planning for Next-Generation SANs

with 4 Gbps Fibre Channel Switches from Brocade

By the end of 2005, most major switch, router, host bus adapter (HBA), and storage vendors are expected to offer 4 Gbps products in key markets. The introduction of 4 Gbps storage area network (SAN) products—such as the Brocade® SilkWorm® 4100 fabric switch, high-speed tape systems, and HBAs—will enable organizations to begin deploying next-generation SAN environments. This article provides guidelines for introducing 4 Gbps devices into enterprise SAN environments based on the features of the SilkWorm 4100 switch.

BY SPENCER SELLS

Related Categories:

Brocade

Dell/EMC storage

Fibre Channel

Fibre Channel switches

Planning

Storage

Storage architecture

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

In many high-speed IT environments, servers and storage devices have already saturated existing 2 Gbps interfaces. New 4 Gbps devices are being designed particularly to help satisfy the needs of high-performance, I/O-intensive applications and burgeoning storage area network (SAN) infrastructures. Such technical advances can lead to business benefits as well. For example, 4 Gbps Fibre Channel technology can help reduce the amount of time required for data backups compared to 2 Gbps Fibre Channel technology, which in turn helps minimize the need to purchase additional servers and storage devices. In this way, 4 Gbps technology can enable organizations to comply cost-effectively with business regulations that require them to manage and store an increasing amount of data.

From a business perspective, 4 Gbps Fibre Channel switches can provide a higher return on investment than 2 Gbps Fibre Channel switches. For example, one 4 Gbps Inter-Switch Link (ISL) can typically support the same amount of traffic as two 2 Gbps ISLs, which means that a smaller number of 4 Gbps ISLs—and SAN ports—can handle the same amount of data traffic as a larger number of 2 Gbps ISLs. Consequently, SAN ports that are no longer needed for ISLs can be freed to connect additional servers and storage. Thus, 4 Gbps switches can also enhance the overall value of the SAN investment compared to 2 Gbps switches.

Because 4 Gbps Fibre Channel technology is based on the same standards as 1 Gbps and 2 Gbps Fibre Channel technology, vendors are expected to introduce 4 Gbps products at approximately the same price point as comparable 2 Gbps products. This will enable organizations to continue to use their existing 1 Gbps and 2 Gbps devices and management tools as they phase in 4 Gbps technology to serve evolving business requirements.

Introducing the Brocade SilkWorm 4100 switch

As the industry's first generally available 4 Gbps SAN switch, the Brocade SilkWorm 4100 switch includes several features and capabilities that can influence how organizations design their SAN environments. With all ports auto-sensing for 1 Gbps, 2 Gbps, and 4 Gbps operations, the SilkWorm 4100 is designed to integrate seamlessly with existing 1 Gbps and 2 Gbps systems—enabling deployment in existing SAN infrastructures. At the same time, the SilkWorm 4100 provides enhanced design flexibility with Ports On Demand software that enables administrators to nondisruptively scale the switch from 16 ports to 24 or 32 ports.

To help ensure high performance, the SilkWorm 4100 application-specific integrated circuit (ASIC) provides nonblocking Fibre Channel operations—meaning that all ports are designed to run at 4 Gbps full-duplex in any configuration. Another key performance enhancement is the advanced buffer-to-buffer credits shared among as many as 32 ports—supporting SAN configurations reaching up to 500 kilometers (310.5 miles). Additional advances include enhanced ISL Trunking capabilities, such as 32 Gbps frame-level trunks (with up to eight 4 Gbps links in a single trunk) and Dynamic Path Selection (DPS) for improved load balancing between trunk groups. The following sections describe how organizations can incorporate these capabilities as they design a new SAN environment or expand an existing SAN with 4 Gbps devices.

Designing SAN environments for 4 Gbps devices

When designing SANs with 4 Gbps devices such as the SilkWorm 4100 Fibre Channel switch, system architects should first determine where and how they will deploy 4 Gbps devices within the SAN environment. One common approach is to use 4 Gbps technology in ISLs to connect switches in an extremely high-performance fabric. Higher per-port bandwidth enables system architects to use fewer ports as ISLs to achieve the same performance level, thereby freeing up ports for servers and storage devices.

New 4 Gbps devices are being designed particularly to help satisfy the needs of high-performance, I/O-intensive applications and burgeoning SAN infrastructures.

This year, early adopters are expected to deploy 4 Gbps devices such as the SilkWorm 4100 switch in stand-alone configurations, as either the core or edge of smaller core-to-edge networks. In an existing 2 Gbps SAN, administrators are likely to deploy devices such as the SilkWorm 4100 switch at the edge of a core-to-edge fabric. Therefore, the 4 Gbps switch would be connected to the existing core switches that have 2 Gbps interfaces. In this case, administrators would follow the existing ISL over-subscription ratio because the ISLs would run at 2 Gbps and all the other devices in the fabric would be 2 Gbps.

As soon as organizations begin deploying 4 Gbps devices, however, administrators may consider adding ISLs or upgrading the core switches to 4 Gbps. *Note:* Administrators can use tools such as Brocade Advanced Performance Monitoring software or the Brocade SAN Health utility to collect ISL usage statistics and determine whether this step is necessary.

When 4 Gbps Fibre Channel blades become available for use in SAN directors such as the Brocade SilkWorm 24000, any 2 Gbps directors at the edge of the fabric will likely receive blades with 4 Gbps capabilities. Likewise, organizations probably will upgrade the blades within directors at the core of large SANs (4 Gbps Fibre Channel blades purchased and old blades transferred to directors on the edge). Or in some cases, administrators may move the entire core chassis to the edge of the fabric where it can support applications that do not require the highest levels of performance.

If deploying the SilkWorm 4100 switch in an exclusively 4 Gbps SAN fabric, if designating some part of the fabric as exclusively 4 Gbps, or if using the SilkWorm 4100 switch for its enhanced distance extension capabilities, administrators should consider the following configuration methods:

- **Long-distance trunking:** If the SAN extends over dark fiber, coarse wavelength division multiplexing (CWDM), dense wavelength division multiplexing (DWDM), or Synchronous Optical Network (SONET), the SilkWorm 4100 switch is designed to provide capabilities that significantly enhance performance over long distances. For example, because the SilkWorm 4100 switch supports up to 255 buffer credits on a single port, it can enable full-speed 1 Gbps operations at approximately 500 kilometers (310.5 miles), 2 Gbps operations at approximately 250 kilometers (155.3 miles), and 4 Gbps operations at approximately 125 kilometers (77.6 miles).
- **Link balancing:** If the SilkWorm 4100 switch is part of an exclusively 4 Gbps network or if a certain portion of the network uses only 4 Gbps ISLs or trunks, administrators should consider reducing the total number of ISLs by half, compared to how many ISLs they would typically use in 2 Gbps fabrics. (As always, a minimum of two ISLs between any edge switch and the core is required to help ensure high availability.) If the fabric has a mix of 2 Gbps and 4 Gbps speeds, the number of ISLs would likely fall somewhere between what would be required for exclusively 4 Gbps and for exclusively 2 Gbps speeds, depending on traffic patterns. Note that SilkWorm 4100 switches support eight-way trunking and DPS between trunks. In many cases, these two features can

Ports trunked	4 Gbps		2 Gbps		1 Gbps	
	Throughput	Distance	Throughput	Distance	Throughput	Distance
8	32 Gbps	30 km/ 18.6 mi	16 Gbps	60 km/ 37.3 mi	Not applicable	Not applicable
4	16 Gbps	60 km/ 37.3 mi	8 Gbps	125 km/ 77.6 mi	Not applicable	Not applicable
2	8 Gbps	125 km/ 77.6 mi	4 Gbps	175 km/ 108.7 mi	Not applicable	Not applicable
1	4 Gbps	125 km/ 77.6 mi	2 Gbps	250 km/ 155.3 mi	1 Gbps	500 km/ 310.7 mi

Figure 1. Extended trunking distances and data transfer speeds

help improve performance over 2 Gbps links as much as upgrading to 4 Gbps interfaces alone improves performance, and DPS can operate even when the SilkWorm 4100 switch is connected to 2 Gbps switches.

Implementing long-distance trunking

The SilkWorm 4100 switch has a total of 1,024 buffer credits shared among as many as 32 ports to support increased trunking distances in extended SAN environments. Of these 1,024 buffer credits, 24 are used for the embedded port; the remaining credits are available for administrators to use.

F_ and FL_Ports receive eight buffer credits by default, and local E_Ports (L0 mode) receive 26 buffer credits (the same amount of credits as 2 Gbps ports). A minimum of eight buffer credits is reserved for each port to help ensure that no ports are starved of credits. The remaining credits are available in a buffer pool, which administrators can configure for use by any of the 32 ports.

In contrast to earlier-generation Brocade switches, the SilkWorm 4100 switch does not automatically assign buffer credits from the pool because line speed can easily be achieved with eight buffer credits for a local device. If administrators need to assign additional buffer credits to a specific port, they can easily do so with the `portCfgLongDistance` command as long as they have a valid Brocade Extended Fabrics software license.

As a result, a single port can have up to 255 buffer credits, providing the enhanced distance capabilities shown in Figure 1. To achieve these longer distances, administrators must deploy a SilkWorm 4100 switch (or 4 Gbps Fibre Channel blades for the SilkWorm 24000 Director, when they become available) on both ends of the long-distance link.

Balancing SAN traffic across switches

Even in over-provisioned networks, administrators may detect “hot spots” of congestion—some data paths running at their limit while others remain unused. In such cases, the network can experience a performance bottleneck even if it has sufficient bandwidth to deliver all data flows without constraint. Brocade switches offer three options to help balance data traffic while increasing both availability and performance:

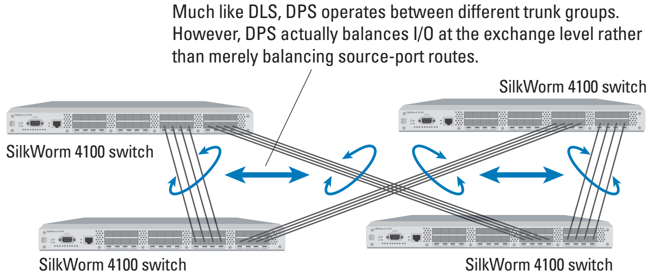
- Source-port route balancing via Fabric Shortest Path First (FSPF), also known as Dynamic Load Sharing (DLS)
- Frame-level trunking between ASICs, also known as Advanced ISL Trunking
- Exchange-level trunking, also known as DPS

Unique to the SilkWorm 4100 switch, exchange-based DPS can optimize fabric-wide performance and load balancing by automatically routing data to the most efficient available path in the fabric. Exchange-based DPS augments ISL Trunking to enhance load balancing in certain configurations, such as routing data between multiple trunk groups. As a result, a combination of DPS and ISL Trunking can provide excellent design flexibility and optimal load balancing.

How DPS works

DPS works by striping Fibre Channel exchanges across equal-cost paths—that is, the sender places an exchange ID into every Fibre Channel frame header. Under normal operation, the exchange ID remains consistent for the duration of a SCSI operation. When a DPS-enabled platform receives a frame, it takes into account all equal-cost routes and calculates the egress port from that set based on a formula using the sender ID (SID), destination ID (DID), and exchange ID (OXID). The formula always selects the same path for a given SID-DID-OXID set.

For most Fibre Channel devices, DPS stripes I/O at the SCSI level. For a given “conversation” between a server and a storage port, one SCSI command would travel the first path, and the next command would travel a different path. All frames within a given exchange would be delivered in order by virtue of traveling the same


Much like DLS, DPS operates between different trunk groups. However, DPS actually balances I/O at the exchange level rather than merely balancing source-port routes.
SilkWorm 4100 switch
SilkWorm 4100 switch
SilkWorm 4100 switch
SilkWorm 4100 switch

Frame-level trunks between SilkWorm 4100 switches can be up to eight links wide, with each link running at 4 Gbps. This example shows four 4 Gbps links trunked into a 16 Gbps pipe between each edge and each of two cores. DPS balances the two trunks into a single 32 Gbps path between edge switches.

Figure 2. Using frame-level trunking plus DPS to load balance across switches in a core-to-edge network

network path. Consider the following: If two servers are writing to two different storage ports across the same network, in-order delivery between the different servers is not important. It only matters that in-order delivery occurs within the data stream sent by each server, not *between* the two different and unrelated data streams.

The technical advantages of DPS

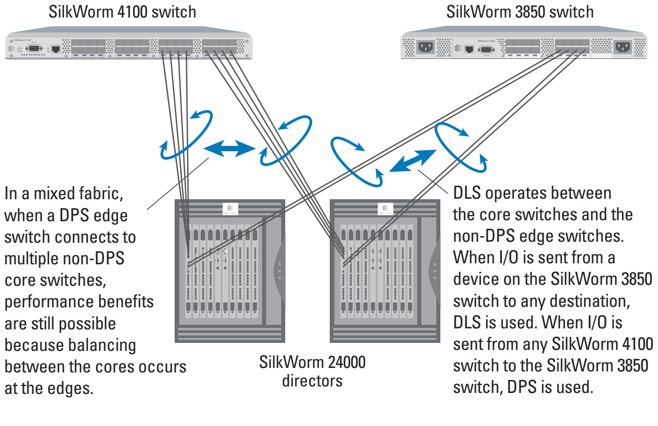
Based on exchange-level trunking, the DPS feature at the heart of the SilkWorm 4100 switch is designed to offer high performance for a SAN. Moreover, DPS can be combined with frame-level trunking to help administrators achieve maximum performance and availability along with a variety of other benefits.

First, DPS does not need to occur within ASIC port groups the way frame-level trunking does. This enables load balancing across different core switches in a core-to-edge network (see Figure 2) or across different blades within a director, rather than simple DLS-based route balancing.

Administrators can balance several groups of ports by using frame-level trunking, and then balance the resulting trunk groups by using exchange-based DPS. This approach can provide the optimal balance of performance (frame-level trunking is faster) and availability (exchange-level trunking balances high-availability network topologies).

In addition, DPS can balance I/O sent from a DPS-enabled platform to any other platform, even if the destination does not support the DPS feature. The transmitting switch selects the path, and the receiver does not need to do anything to ensure in-order delivery. This approach enables full backward compatibility with existing switches as well as performance benefits—even if not all of the switches in a fabric are using the latest technology (see Figure 3).

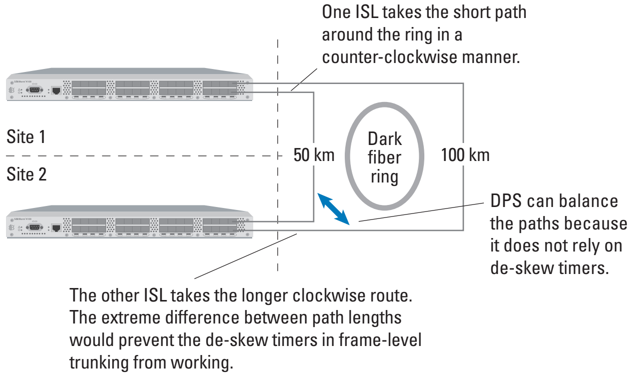
DPS can also balance I/O across long-distance configurations not supported by frame-level trunking. For example, if two SAN sites have two links that take substantially different paths, too much skew might prevent the formation of a frame-level trunk. In contrast, DPS would be able to balance these links, because it does not rely on de-skew timers for in-order delivery (see Figure 4).



In a mixed fabric, when a DPS edge switch connects to multiple non-DPS core switches, performance benefits are still possible because balancing between the cores occurs at the edges.

DLS operates between the core switches and the non-DPS edge switches. When I/O is sent from a device on the SilkWorm 3850 switch to any destination, DLS is used. When I/O is sent from any SilkWorm 4100 switch to the SilkWorm 3850 switch, DPS is used.

Figure 3. Using DPS in a mixed-speed Fibre Channel fabric



One ISL takes the short path around the ring in a counter-clockwise manner.

DPS can balance the paths because it does not rely on de-skew timers.

The other ISL takes the longer clockwise route. The extreme difference between path lengths would prevent the de-skew timers in frame-level trunking from working.

Figure 4. Using DPS in large fiber ring configurations

When considering how to utilize these options, administrators should note that any SAN with more than one ISL can benefit from link balancing to help ensure high performance. The most effective designs combine frame-level trunking for high performance, and DLS or DPS to balance multiple trunk groups for high availability.

Migrating to 4 Gbps switch technology

With the introduction of the Brocade SilkWorm 4100 switch—the industry's first 4 Gbps Fibre Channel SAN switch—organizations can begin implementing high-performance SANs based on next-generation technology. In addition to providing the performance required for data-intensive applications, these 4 Gbps devices can help reduce deployment costs because SAN designs typically require fewer 4 Gbps ISLs (and consequently less management overhead) than 2 Gbps switches to achieve the same level of performance. Moreover, backward compatibility with existing 1 Gbps and 2 Gbps SAN environments can help protect current technology investments and provide excellent return on investment for years to come.

Spencer Sells is the director of product management at Brocade. He has spent 11 years in the high-tech industry, working with mainframe servers, enterprise storage, midrange storage, and storage networking at companies such as Brocade, Amdahl, and Gadzoox.

FOR MORE INFORMATION

Brocade SilkWorm 4100 switch:
www.brocade.com/products/switches/silkworm_4100

Brocade Conference 2005
www.brocade.com/conference2005

Make the Connection **with Brocade**

Our experience will
help you meet your objectives

Since 1995, Brocade® has been a pioneer in defining the networking model for storage. Today, we lead the industry in delivering innovative SAN solutions that improve application performance and reduce your overall storage costs.

To make the connection with your business, visit www.brocade.com.

A Comprehensive Family of SAN Infrastructure Solutions

Switches Directors Routers Embedded Server Blades Software Technical Resources

BROCADE

Efficient, Cost-Effective

Data Protection for Remote Offices

Using VERITAS Backup Exec and VERITAS Replication Exec on Dell Hardware

Distributed enterprises require cost-efficient methods to protect ever-increasing amounts of critical business data accumulating at multiple office locations. By using VERITAS® Backup Exec™ and VERITAS Replication Exec™ software running on Dell hardware, IT organizations can consolidate backups from branch offices, helping reduce operational expenses and enhance IT staff productivity. This article discusses how one financial services organization, Farmers & Merchants Bank, implemented VERITAS software on a Dell™ PowerVault™ 775N network attached storage (NAS) server to enable centralized, automated backups for each of its 20 branches.

BY BRIAN NELSON AND BRIAN GREENE

Related Categories:

Case study

Data archiving

Dell PowerVault storage

Disaster recovery

Storage

VERITAS

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

Many of today's enterprises generate and rely on tremendous volumes of critical data for their daily business interactions. IT organizations in such enterprises are tasked with minimizing the risk of loss to mission-critical data by implementing efficient backup and storage methods—while keeping the costs associated with data backup and archiving low. However, for organizations with multiple office locations, ensuring consistent, cost-efficient data protection for each site has become a daunting challenge. Many organizations have gone to great lengths to protect business data at each location, deploying staff, tape, tape drives, and backup software to perform local backups at each site. Not only is this approach exhaustively expensive, but it is also increasingly time-consuming and cumbersome for administrators to manage. Distributed enterprises need a solution that will help them protect company-wide data consistently and effectively while keeping IT costs and administration to a minimum.

Financial services institutions are a representative example of the type of enterprise that depends on keeping critical data from multiple office locations highly available for its daily business operations. This article examines how one Southern California financial services institution, Farmers & Merchants Bank (F&M), implemented VERITAS Backup Exec and VERITAS Replication Exec software on a Dell PowerVault 775N network attached storage (NAS) server to provide centralized, automated data storage—helping streamline IT operations and minimize data storage and archiving costs.

Reevaluating a distributed backup infrastructure

F&M is a privately held financial institution that has been serving individual and business customers through its 20 branches in Los Angeles and Orange counties since its founding in 1907. As is the case in many geographically distributed enterprises, individual branches of F&M were responsible for backing up and archiving their own data on a daily basis.

Over time, the process of daily backups at individual branches became increasingly more expensive for F&M as the bank expanded to new locations and grew its

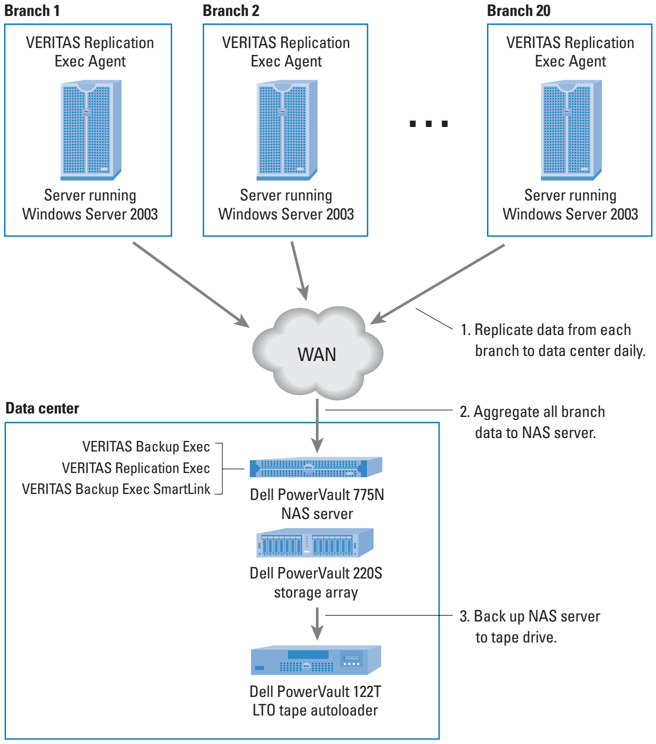

Branch 1
VERITAS Replication Exec Agent
Server running Windows Server 2003
Branch 2
VERITAS Replication Exec Agent
Server running Windows Server 2003
. . .
Branch 20
VERITAS Replication Exec Agent
Server running Windows Server 2003
WAN
1. Replicate data from each branch to data center daily.
Data center
2. Aggregate all branch data to NAS server.
VERITAS Backup Exec
VERITAS Replication Exec
VERITAS Backup Exec SmartLink
Dell PowerVault 775N NAS server
Dell PowerVault 220S storage array
3. Back up NAS server to tape drive.
Dell PowerVault 122T LTO tape autoloader

Figure 1. F&M backup infrastructure enabling centralized, automated backup and restore functions

customer base. The backup window for each branch ballooned from a few hours to as many as eight hours in some cases. Branch backups were often unreliable and had to be repeated, creating a risk of data loss and an administrative headache for the IT team in the bank's centrally located data center, which was responsible for overall data operations. In addition, as branches experienced staff turnover, new administrators at individual locations had to be trained in the bank's complicated backup and archiving procedures, further reducing IT efficiency and driving up costs.

F&M responded to these challenges by migrating to a Microsoft® Windows® 2000 Server OS–based client/server architecture in 1999 and replacing its existing backup software with VERITAS Backup Exec to help improve the speed and reliability of its backups (for more information about VERITAS products, visit www.veritas.com/products). VERITAS Backup Exec demonstrated its effectiveness as a backup and business continuance tool. For example, when a secondary drive on a RAID storage system at a local branch failed before the local staff could initiate repairs on the faulty primary drive, the bank's IT organization used Backup Exec to restore data and get the branch back up and running in only a few hours. However, despite the highly effective backup performance and business continuance capabilities that VERITAS Backup Exec demonstrated in situations like this across the bank's 20 branches, that architecture alone did not fulfill all the bank's cost and efficiency objectives. Prompted by concerns about the expense of performing individual backups at each branch—and the financial and legal liabilities of a dispersed data protection system—F&M sought to convert its distributed storage architecture into a centralized, automated data backup and archiving infrastructure.

Before embarking on the project, the F&M IT staff reexamined the bank's data storage and archiving infrastructure and determined that a centralized backup infrastructure would require the new IT environment to provide the following capabilities:

- **Centralization:** Replicate all important business information daily from branches to a central location to eliminate individual backups at branches
- **Automation:** Automate daily backups to minimize administration
- **Data retention:** Ensure accurate backup and retention of mission-critical data
- **Business continuance:** Restore lost data quickly to help avoid business interruption
- **Scalability:** Scale the storage architecture efficiently and cost-effectively as required to accommodate ever-increasing data volumes

Implementing a centralized backup architecture using VERITAS software on Dell hardware

In November 2004, VERITAS Technical Support engineers, Dell Professional Services, and the F&M IT staff implemented centralized, automated backup and restore functions at the bank's data center using VERITAS software and Dell hardware (see Figure 1). The team aggregated backups from each branch to a NAS environment at the data center, which was powered by a Dell PowerVault 775N NAS server running Microsoft Windows Storage Server 2003 and a Dell PowerVault 220S storage array running Microsoft Windows Server™ 2003, Enterprise Edition—both connected to a Dell PowerVault 122T Linear Tape-Open (LTO) tape autoloader.

The team installed VERITAS Backup Exec 10 software on the PowerVault 775N NAS server in the F&M data center to provide fast disk-based backup and recovery as well as centralized management and monitoring capabilities. VERITAS Replication Exec 3.1, which was also installed on the NAS server in the data center, allowed F&M to dispense with daily backups at individual branches by replicating data from branches to the data center each day. To implement VERITAS Replication Exec at individual branches, the team installed Replication Exec on a central backup server and push-installed this software on a server running Windows Server 2003, Enterprise Edition, at each branch. Replication Exec operates by copying data from each branch-office server to the central data center over an IP-based wide area network (WAN) to enable consolidated backups.

Enabling IT efficiencies through centralized backups

Using VERITAS software on Dell hardware, F&M achieved its goal of centralized, automated backups. Individual branches no longer perform tape backups—greatly minimizing the bank's overall administrative burden. Every evening, VERITAS Replication Exec replicates data from each branch to the NAS server at the F&M data center, where VERITAS Backup Exec writes the complete data set onto tape—helping ensure that backups are highly reliable.

VERITAS software provides several features that help simplify administration for distributed enterprises that want to consolidate and automate their data backups, including centralized management, automated data protection, maximized data storage resources, streamlined data recovery, and cost-effective scalability.

Centralized management. Replication Exec integrates seamlessly with Backup Exec so that administrators can monitor the progress of Replication Exec jobs at all 20 branches from one centralized Backup Exec console. This centralized monitoring capability is made possible by VERITAS Backup Exec SmartLink, a feature recently introduced in Backup Exec 10. SmartLink helps save time by enabling administrators to monitor all replication activities from within the Backup Exec console. From the central Backup Exec console, administrators can log in and track each replication job individually and receive alert notification when a replication job has failed—further easing administrative burdens.

Automated data protection. Replication Exec can be installed, configured, and managed from a single, central console. After being installed on the central backup server, Replication Exec can be push-installed to each remote server. Once replication rules are configured, Replication Exec runs in an automated fashion, helping protect branch-office data with minimal IT staff involvement.

Maximized data storage resources. The flexible scheduling features of Replication Exec enable F&M to back up data continuously or at a scheduled time. In addition, Replication Exec allows system administrators to choose which files to back up—helping to save tape space at the data center and minimizing bandwidth consumption by avoiding unnecessary backups. F&M uses the large drive capacities of the Dell PowerEdge 220S to store replicated data on the NAS server. By doing so, the bank can restore weeks of data without having to retrieve tapes from long-term storage.

Streamlined data recovery. VERITAS software helps simplify a variety of daily data recovery tasks through its drag-and-drop capability and intuitive VERITAS Backup Exec interface. These ease-of-use features help administrators manage backup and restore processes with virtually no training.

Cost-effective scalability. VERITAS software and Dell hardware are designed to offer economical scalability and high storage capacity. Replication Exec scales easily to extend backup protection to a virtual limitless number of branches. This enables administrators to maximize personnel and backup resources at a central office while leveraging the scalability of cost-effective Dell hardware.

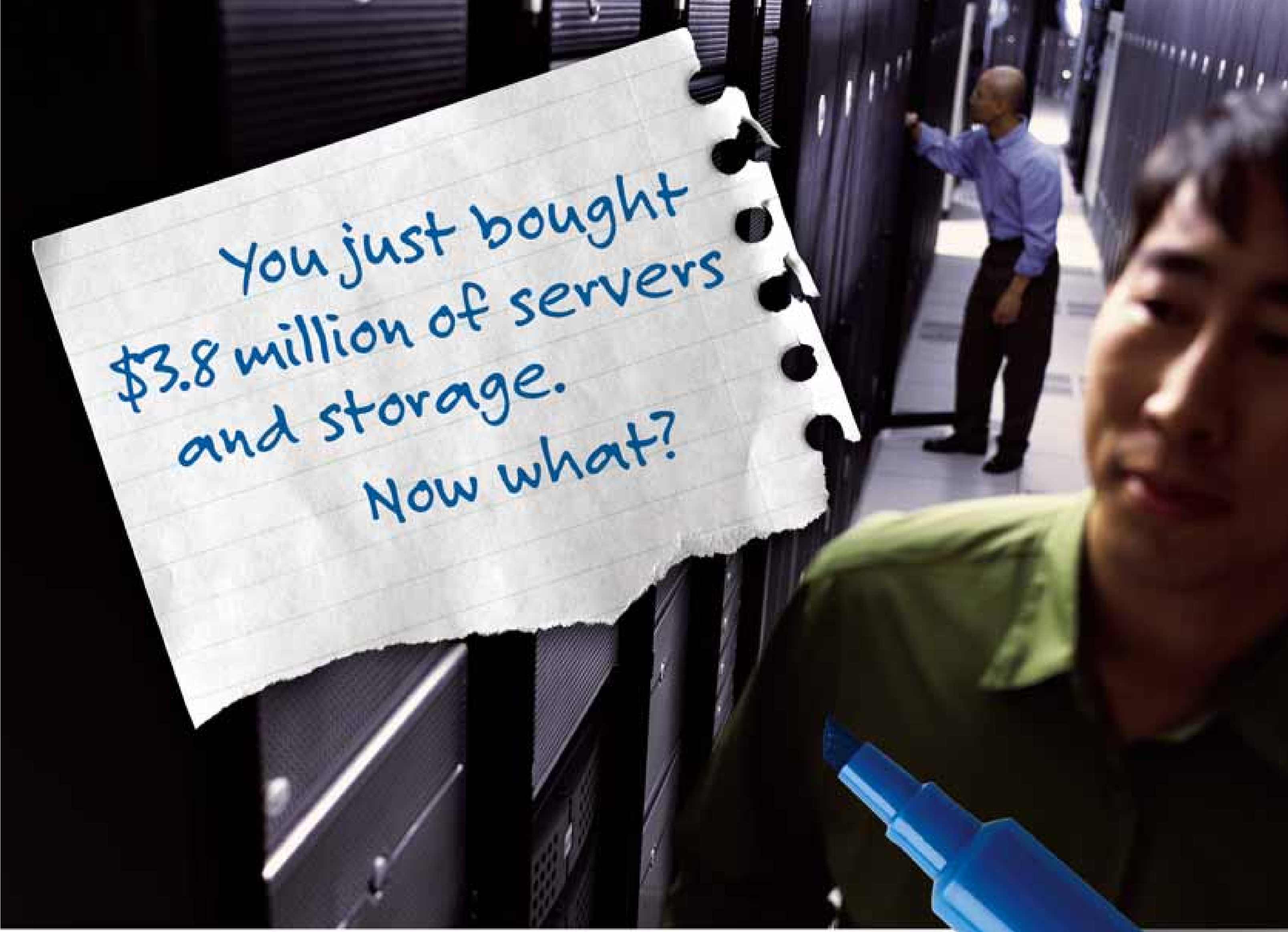
Enabling productivity and performance for future growth

The preceding efficiencies, which are enabled by the high level of integration and performance that VERITAS software and Dell hardware offer in a centralized architecture, can enable significant IT productivity gains—particularly for enterprises coping with growing data storage volumes. For example, F&M Bank estimates that the elimination of tape backups at individual branches has reduced the bank's IT administrative labor requirements by 24 hours per week, which has also resulted in an associated reduction in IT costs.

In addition, the combined backup of daily data from all 20 bank branches now takes only about four hours—significantly less than the eight hours each individual branch used to take. This enhanced efficiency is critical because the bank estimates that its data backup and storage requirements will grow at an annual rate of 25 percent over the next five years. For F&M and other enterprises charged with the challenge of serving an expanding customer base or simply accommodating an increased volume of data to comply with government regulations regarding data retention, VERITAS Backup Exec and VERITAS Replication Exec software together with Dell PowerVault NAS servers can help provide cost-effective performance and functionality and the flexibility to adapt to future needs.

Brian Nelson is a product marketing manager at VERITAS Software Corporation, a leading storage software company that offers products to improve application performance and provide data protection, storage management, high availability, and disaster recovery.

Brian Greene is a senior staff product manager at VERITAS Software Corporation.


You just bought
$3.8 million of servers
and storage.
Now what?

Meeting Regulatory Compliance Requirements with VERITAS Enterprise Vault and Microsoft Windows Server Technologies

Creating an electronic messaging system to meet regulatory compliance requirements can be a complex process. This article provides guidance for organizations that need to implement such a system, outlining a strategy that is designed to enable centralized search and retrieval across archived data as well as workflow processes that can help provide quick response to requests for archived electronic data or correspondence. This approach allows fast, efficient data retrieval that can free administrators from the process of restoring messaging data from backup tapes.

BY SCOTT ROSEN

Related Categories:

Dell PowerEdge servers

Dell PowerVault storage

Microsoft Exchange

Microsoft Windows

Network attached storage (NAS)

Regulatory compliance

Sarbanes-Oxley Act

Storage

VERITAS

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

To meet regulatory compliance requirements, some organizations must preserve e-mail correspondence as a business record that can withstand scrutiny in a court of law or regulatory review. To help address this need, data archiving tools such as Enterprise Vault™ software and Discovery Accelerator from KVS, a business unit of VERITAS, can be combined with Microsoft® Exchange Server 2003 and Dell™ PowerVault™ network attached storage (NAS) servers powered by Microsoft Windows® Storage Server 2003. VERITAS Enterprise Vault software provides an enterprise-class platform for data archiving, and Discovery Accelerator enables centralized search and retrieval across archived data as well as workflow processes that help administrators respond quickly to requests for archived documents and messages. In addition, Enterprise Vault software and Discovery Accelerator can be integrated with Microsoft Windows Storage Server 2003, Microsoft Exchange Server 2003, and Microsoft SharePoint® Portal Server 2003 to create a prescriptive framework for message and document archiving.

The configuration described in this article includes Dell PowerVault NAS servers as well as Dell PowerEdge™ servers designed to provide a high level of performance, scalability, and reliability. Together, these products offer organizations key components of an overall information life-cycle management (ILM) strategy: the ability to capture, archive, and destroy data based on corporate policies; an audit trail that enables compliance; and dependable tools that leverage existing IT investments.

This article discusses the use of message archival systems and explores protection strategies that can help ensure that enterprise data is appropriately stored and maintained. Unless otherwise noted, the approach outlined in this article assumes that Microsoft Windows Server™ 2003, the Microsoft Active Directory® directory service, and Exchange Server 2003 have been deployed in the data center.

Understanding the current business environment and regulatory compliance

Over the past decade, e-mail has become a mission-critical tool for many enterprises. However, e-mail archive and retrieval procedures are usually enacted in an ad hoc manner. Few organizations take the time to clearly define policies regarding the use of messaging, the types of data that will be transmitted, and the types of data protection to use. Many organizations are discovering the need for a system that can help ensure that data within their Exchange Server messaging environments is safely stored in a searchable, retrievable format.

Although many regulations affecting businesses do not necessarily require message archiving, today's regulatory environment is changing and businesses need to be aware of the influence this change may have on the long-term operations of their messaging systems. Businesses in the financial and health-care industries have long been aware of the need to archive and track their communications because of regulations such as the Securities and Exchange Commission (SEC) Rule 17A-4 and the Health Insurance Portability and Accountability Act of 1996 (HIPAA). Now, organizations in industries that have not previously felt the need to retain e-mail data may face that necessity. Regulations such as the Sarbanes-Oxley Act of 2002 have highlighted the need for organizations in all industries to maintain, store, and secure data, including electronic (instant) messages, for periods ranging from 90 days to 30 years or more.

Identifying the components of a message archiving system

The approach outlined in this article incorporates Microsoft Exchange Server 2003, Microsoft SharePoint Portal Server 2003, and Microsoft Windows Server 2003 running on Dell PowerEdge servers to provide messaging services. VERITAS Enterprise Vault software and Discovery Accelerator also run on Dell PowerEdge servers, and these applications archive information to Microsoft Windows Storage Server 2003, which runs on Dell PowerVault NAS servers. Figure 1 shows the overall architecture of this message archiving system.

Microsoft Exchange Server 2003

Message journaling within Exchange Server enables organizations to archive messages sent between end users and external Internet addresses. With minor configuration changes, all internal messages can also be archived. Service Pack 1 for Exchange Server 2003 introduces enhanced journaling capabilities that enable VERITAS Enterprise Vault software to provide a rich archiving tool set to organizations.

VERITAS Enterprise Vault and Business Accelerators

Delivering enterprise-class document and e-mail archiving services for Microsoft Exchange and SharePoint Portal Server implementations, Enterprise Vault 5 with Cumulative Patches 3 offers a single interface for archived e-mail messages, SharePoint file system documents, and instant messages. Enterprise Vault software enables administrators to consolidate Exchange servers, eliminate Personal Folders (.pst) files from the environment, archive data within file servers, migrate data within Exchange public folders, archive current data within mailboxes, and meet regulatory compliance goals.

The content archiving approach provided by Enterprise Vault, along with the Discovery Accelerator add-on, can help reduce the ongoing cost of e-mail storage, bring control to mailbox management, optimize the backup and recovery cycle, and ensure that valuable information can be retrieved quickly and efficiently to facilitate compliance and knowledge management. Also, Compliance Accelerator for Enterprise Vault can be implemented to provide additional capabilities that help meet ILM requirements and alleviate business risk.

Microsoft Windows Storage Server 2003

In the approach described in this article, Windows Storage Server 2003 is configured to host the Enterprise Vault archives. Windows Storage Server 2003 is designed to provide dependability and seamless integration while enabling organizations to derive optimal value from their networked storage. For example, Windows Storage Server 2003 is well suited for consolidating organizational data such as Enterprise Vault archives into a single system that can help achieve cost reduction and policy-based management of storage resources.

Windows Storage Server 2003 includes advanced availability features such as point-in-time data copies, replication, and multi-node clustering. Because Windows Storage Server 2003 implementations

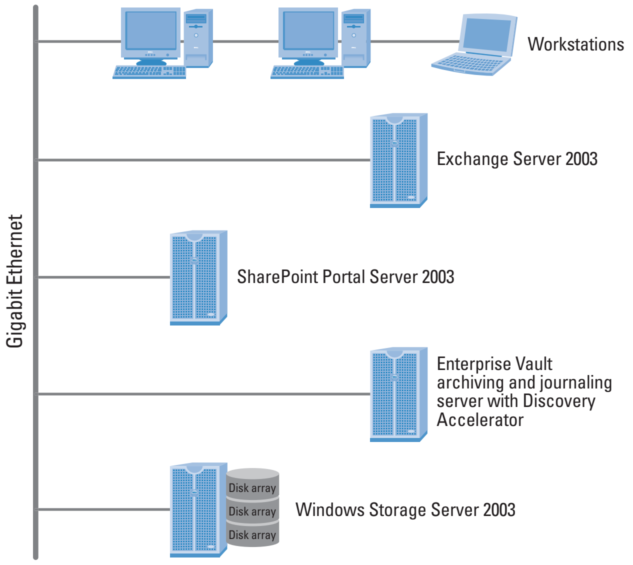

Workstations
Exchange Server 2003
SharePoint Portal Server 2003
Enterprise Vault archiving and journaling server with Discovery Accelerator
Disk array
Disk array
Disk array
Windows Storage Server 2003
Gigabit Ethernet

Figure 1. Components of the message archiving architecture

are typically preconfigured to the purchaser's specifications, they can be rapidly deployed out-of-the-box and require minimal expertise to set up. The Web-based user interface makes management easy. The Dell PowerVault 700 series storage servers can be preconfigured with Windows Storage Server 2003.

Defining data retention policies

Date retention is both an IT and a business concern that is defined by multiple groups within an organization, including the legal, IT, finance, and operations departments. A sufficient electronic messaging policy defines acceptable use of the system, such as permitting end users to send and receive personal e-mail, allowing e-mail solicitations, disallowing the use of e-mail for harassing or threatening messages, and prohibiting the transmission of potentially offensive images. The policy should define what company materials are confidential, and when and under what circumstances company-confidential materials can be shared with third parties. It is preferable to clearly state that users cannot send company-confidential data to a third party unless that third party is receiving the data for a legitimate business reason, and that illegal use of the system will not be tolerated. Retention periods for communications should be clearly defined; organizations that are subject to specific regulations defining retention periods should ensure that these requirements are clearly stated in their policies.

When creating an electronic messaging policy, organizations should ensure that the correct stakeholders are involved and that the policy is not created in a vacuum. The legal department, financial advisors, and systems managers must coordinate their efforts to create a policy that not only is legally correct, but also adequately protects the interests of the organization and can be properly implemented and enforced. The risks and realities of the organization's structure must be considered, and the policy should be clearly defined and implemented.

The SANS (SysAdmin, Audit, Network, Security) Institute, a cooperative research and educational organization for information security professionals, provides a sample policy for e-mail retention, available at www.sans.org/resources/policies/e-mail_retention.pdf. This can help organizations begin the process of creating a data retention policy.

Examining the current regulatory environment

In the United States, numerous federal regulations affect various organizations.[1] While the financial industry has long been subject to oversight by the SEC and the National Association of Securities Dealers (NASD), and the health-care industry has rushed to meet the requirements put in place by HIPAA, other types of organizations are now becoming actively involved in the regulatory process. The enactment of broad-reaching regulations, such as the Gramm-Leach-Bliley Act and the Sarbanes-Oxley Act, has created the need for organizations in various industries to identify ways to safeguard, disseminate, store, and track financial information. Many states have enacted regulations that supersede these federal regulations, so organizations also must ensure that they are complying with the pertinent state laws in addition to applicable federal regulations.

Regulations affecting electronic messaging

Regulations can affect how, where, and how long organizations must maintain electronic records, including e-mail. Compliance with the relevant regulations is a complex process and should be overseen by appropriate legal counsel. While the following regulations are pertinent to many organizations and present a good overview of the overall regulatory environment today, organizations should rely on legal counsel to determine applicability and analysis:

- Sarbanes-Oxley Act
- SEC Rule 17A-4
- Gramm-Leach-Bliley Act (including the Financial Institution Privacy Protection Act of 2001 and Financial Institution Privacy Protection Act of 2003 amendments)
- HIPAA
- Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act of 2001 (Patriot Act)
- Department of Defense Rule 5015.2-STD
- National Archives and Records Administration
- CFR Title 47, Part 42 – Telecommunications
- CFR Title 21, Part 11 – Pharmaceuticals

A more detailed discussion of these regulations can be found at www.veritas.com/compliance/home.html.

Retaining and archiving messaging data

While many regulations require a specific retention period for business-specific data, every business is not required to meet this requirement. Financial services organizations typically have the most stringent data retention requirements. Nonetheless, enterprises that are not subject to specific data retention requirements should document their data retention policies and follow these policies. Otherwise, organizations that define but do not follow their data retention policies may be required to spend innumerable hours restoring and retrieving data from backup media during a legal discovery process.

A well-thought-out data management plan should ensure that data retention policies mesh with actual data management processes.

[1] The regulations referenced in this article are specific to the United States, but many other countries have similar legislation in place. Organizations should be aware of the regulatory requirements for the geographic areas in which they conduct business.

For example, if the data management plan states that e-mail is kept for one year, backup tapes should not be retained for more than that length of time. Centralized data storage for e-mail and other types of documents—such as a SharePoint Portal Server document repository—can help administrators recover such data easily. The ability to preview this documentation, should it be necessary, can be extremely helpful when approaching a legal proceeding.

Numerous factors can create the need for a regulatory compliance system. Determining whether an organization needs to implement such a system requires cooperation between various divisions of the organization. Understanding the drivers for the implementation of document life-cycle management tools such as Enterprise Vault and Discovery Accelerator can help organizations ensure that the data management plan will have proper system support.

The three primary components of a regulatory compliance system are archiving, retention, and discovery. These components complement one another, and organizations that must meet regulatory compliance rules should consider the factors driving each component before implementing such a system.

Business-critical data should be maintained in a logical, retrievable manner. The challenge of message archiving is determining what data to keep, how long to keep it, who should have access to it (or to a subset of the data), and where to store it.

What data should be archived?

Data that pertains to legal, financial, and business decisions should be archived according to the organization's data retention policy. E-mail messages relating to lunch dates, personal conversations, and the minutiae of running a business probably do not need to be maintained.

Accurate data archiving, with an audit trail, is required to help ensure that all business data is accurately captured and can be verified as original data, or an accurate reproduction thereof. The right data must be captured and stored, and it also must be retrievable.

For the purposes of the data management strategy enabled by the configuration described in this article, the archived data is primarily messaging data. When envelope message journaling is enabled, all messages from, to, and within an Exchange environment are sent to a central journaling mailbox. The data sent to the journaling mailbox is then queued for delivery to the Enterprise Vault server.

How long should data be retained?

Businesses that are bound by the SEC should retain data for no less than seven years; during the first two years of retention, the data must be easily accessible. Other types of industries may have specific regulations that pertain to record keeping, and administrators must understand these regulations and their system requirements. Even businesses that are not bound by industry-specific legislation or rulings are well advised to define a specific data retention period and enact technical measures to comply with that decision.

Of course, data retention policies must be in place before organizations initiate any compliance efforts. It is important that all stakeholders within an organization are involved in creating this policy; IT cannot create a policy without input from other departments.

Who can access the data?

To track communications and help ensure that end users are complying with pertinent regulations, administrators may need to give specific trusted individuals access to stored data. This access should be controlled and audited to prevent abuse. Enterprise Vault software with Discovery Accelerator enables administrators to assign roles to users, thereby helping control data access and retrieval. Authorized reviewers can quickly target and mark specific data as necessary to support legal discovery, compliance-related audits, or investigations. Discovery Accelerator provides a structure to control which users may access data and how data is reviewed.

Who should manage the data?

Few enterprises employ corporate librarians to manage their company's data. Even when this role exists within an organization, it is often underfunded, under-recognized, and under-supported. Organizations that do not have a corporate librarian need tools that can support this function—Microsoft and KVS are teaming up to help fill this role. Organizations that do have corporate librarians can also benefit from implementing the Microsoft-KVS system described in this article, because Enterprise Vault is designed to provide a single point of reference for data that originates from several sources.

Because documents from numerous sources—such as Exchange databases, file servers, and SharePoint Portal Server sites—can be merged into a single Enterprise Vault system, this system is designed to become the authoritative source for information gathering. Knowledge management teams can use this data repository not only for discovery purposes, but also for the purpose of gaining an enhanced understanding of the business value of the data. Corporate librarians and knowledge management teams do more than just find information—they analyze and evaluate data to maximize the utility of the information. Data is an important business asset, and the knowledge management team consists of people who understand the vital nature of an organization's data.

Discovery Accelerator provides roles for data management and retrieval (see Figure 2). These roles include the System Administrator, who creates new cases (discovery processes), configures the marking scheme so messages can be accurately labeled once discovered, and creates user roles; the Case Administrator, who manages the case itself, assigns items to reviewers, and configures new marking schemes; and the Reviewer, who examines the data and marks it for further action, if necessary.

Where should the data be stored?

A centralized data repository can make the discovery process more efficient and reliable than is possible in a widely disparate storage system. Centralized archiving also can be less expensive than distributed storage because it helps provide a better economy of scale for the storage hardware. Although organizations can use either centralized or distributed archiving, most opt to use a centralized architecture for Enterprise Vault software because its caching is designed to provide reliable access to data over long distances and variable electronic link speeds. All business-related data should be kept on servers, and messaging data should be retained on Microsoft Exchange Server or within an archiving system similar to the one described in this article. Local Exchange .pst files pose a risk because they are not centrally controlled and present an unreliable long-term archival system; these files reside on local workstations that all too often are not backed up on a regular basis.

The archiving system described in this article uses NAS servers running Windows Storage Server 2003 to host the Enterprise Vault data archive. Windows Storage Server 2003 is easy to deploy, uses familiar technology for IT administrators, and can be controlled using a Web interface—enabling organizations to quickly add storage to the enterprise network without the need for intense training. Windows Storage Server 2003 also can host several terabytes of data and is designed to provide dependable storage for organizational data.

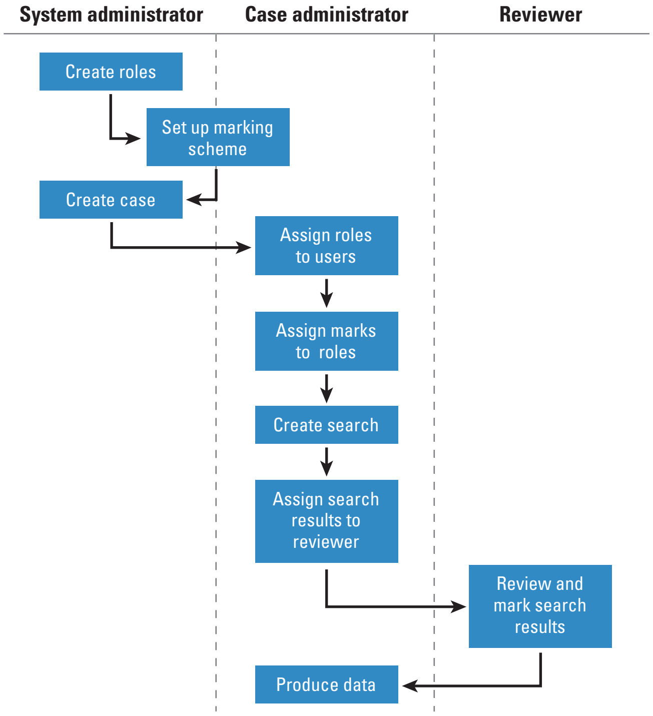


Figure 2. Discovery Accelerator workflow and roles

Reaping the benefits of data retention

Many enterprises that implement a message archiving system similar to the one presented in this article do so to meet regulatory compliance needs. However, implementing such a system can provide additional advantages:

- Archived data creates a searchable corporate knowledge base. Data can be readily searched for and retrieved using tools such as Discovery Accelerator.
- Duplication of effort is reduced. For example, global marking of data within Discovery Accelerator helps ensure that data that has been through a discovery process once does not need to be re-discovered and re-reviewed. This can be particularly helpful when the scope of multiple discoveries overlaps.
- Records of communications and processes that were previously stored in individual mailboxes can be made available to individuals or groups that were not involved in the initial communication or document approval path. This capability allows new employees to understand the history behind past business decisions.
- Archived data helps provide improved business continuity, enabling documentation and communications essential to the long-term success of an organization to be accessed in a single repository that can easily be searched and from which data can be quickly retrieved.
- Archived data helps increase end-user productivity. Because of mailbox size limits, users often resort to storing e-mail messages in .pst files, leading to the need to search multiple sources for an important message or document. Enterprise Vault helps eliminate the need for these files, and it enables information workers to quickly retrieve data using simple searches when the client tools are installed on their workstations or the Web retrieval tool has been made available.
- The ability to verify communications helps mitigate risk to business data, and document life-cycle management enables organizations to oversee data intelligently.

Implementing the message archiving system

The electronic messaging system described in this article is designed for an organization with fewer than 3,000 users that seeks to quickly implement a solid platform to meet information life-cycle needs and regulatory compliance requirements. Enterprise Vault software from VERITAS, together with offerings from Microsoft and Dell, is designed to provide reliable data archiving and to enable organizations to implement a message archiving system with minimal planning.

During the fourth quarter of 2003, Microsoft and KVS engineers demonstrated this approach at Microsoft's labs using the systems described in this article, including Dell PowerEdge servers and Enterprise Vault software. The engineers found that the system architecture and implementation described in this article can provide a good fit for small to medium enterprises and that Windows Storage Server 2003 is well suited to store the Enterprise Vault archives. Microsoft Exchange Server 2003 and SharePoint Portal Server 2003 can be integrated with Enterprise Vault, and the products comprising the integrated architecture described in this article are designed to coexist without requiring numerous custom configuration steps.

Satisfying the growing need for reliable data retention

Organizations that do not have a long-term message archiving system in place should be planning one. To begin this process, administrators must thoroughly understand their system's messaging capabilities, the processes and technologies involved, and the requirements to ensure that clear data retention policies are in place and being followed.

Microsoft Exchange Server 2003, Microsoft SharePoint Portal Server 2003, and VERITAS Enterprise Vault software with Discovery Accelerator can provide a solid foundation to help organizations meet regulatory compliance requirements. In addition, the flexible Dell PowerVault NAS servers and Microsoft Windows Storage Server 2003 enable organizations to quickly and easily implement large disk arrays that can support the storage needs of Enterprise Vault.

Scott Rosen manages the Global Dell Relationship and Appliance Systems for KVS, a business unit of VERITAS Software. His focus is on channel development. He graduated from the University of Michigan with a degree in Organizational Psychology and Finance.

FOR MORE INFORMATION

Enterprise Vault and Discovery Accelerator:
www.kvsinc.com

DELL™

Implementing Cost-Effective RAID

with CERC SATA 2s Disk Technology

Dell recently introduced a low-cost RAID implementation known as the Cost Effective RAID Controller (CERC) SATA 2s, a software-based RAID controller designed for use with Serial ATA (SATA) disk technology. This article examines the architecture, configuration, and management of CERC SATA 2s, including an overview of how CERC SATA 2s interacts with the server's BIOS and how it compares with other RAID implementations.

BY HARISH JAYAKUMAR AND ED MATTHEWS

Related Categories:

Dell OpenManage

Dell PowerEdge servers

RAID

Storage

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

A mainstay for organizations deploying servers, RAID techniques are primarily used to provide protection from disk failures. This protection is achieved by storing data redundantly on multiple disks; if one disk fails, the other disks can still service incoming I/O requests. The various RAID levels represent the different algorithms used to provide data redundancy. Each level has its own cost and performance advantages.

Figure 1 shows a theoretical calculation[1] for the maximum throughput per dollar relative to RAID-0 for different types of RAID.[2] For the purposes of this example, the cost of the RAID hardware is directly proportional to the number of disks the RAID system uses in the disk array. This comparison shows that, given equivalently priced RAID-0 and RAID-1 systems with comparable performance, a RAID-1 system is expected to sustain half the number of small writes per second that a RAID-0 system is expected to sustain. Likewise, a RAID-5 system is expected to sustain a maximum of only one-fourth to one-*N*th (where *N* is the number of disks) the number of small writes per second that a RAID-0 system is expected to sustain.

To limit the CPU burden placed on a system, the Dell™ Cost Effective RAID Controller (CERC) SATA 2s implements only RAID-0 and RAID-1.

Introduction to SATA and CERC SATA 2s

Serial ATA (SATA) disk drives are based on a low-cost technology that is replacing Parallel ATA (PATA) disk drives in low-end servers. SATA implements a serial data transfer mechanism, and the serial interface is designed to provide higher throughput rates than PATA technology.

CERC SATA 2s supports two SATA disk drives and is a driver-based software RAID implementation.[3] To understand

[1] The source of this calculation is "The RAID Tutorial: Cost & Performance Issues" by Israel Koren, Department of Electrical and Computer Engineering, University of Massachusetts, Amherst; for more information, visit www.ecs.umass.edu/ece/koren/architecture/Raid/cp.html.

[2] RAID-0 is actually a misnomer because it does not provide redundancy or protection from disk failures. In a RAID-0 implementation, the data is stored across multiple disks. The RAID-0 approach enables increased performance because multiple drives can concurrently service an I/O request.

[3] CERC SATA 2s is also referred to as CERC SATA 1.5/2s—the number *1.5* stands for the interface speed.

RAID level	Number of disks	Capacity (GB)	Throughput per dollar relative to RAID-0				Relative storage efficiency
			Small read	Small write	Large read	Large write	
RAID-0	2, 3, or 4	*S***N*	1	1	1	1	1
RAID-1	2	*S***N*/2	1	½	1	½	½
RAID-5	3, 4, or 5	*S**(*N*–1)	1	Max (1/*N*, ¼)	1	(*N*–1)/*N*	(*N*–1/*N*)

S: Size of smallest disk
N: Number of disks
Small read or write: An I/O request of one striping unit
Large read or write: An I/O request of one full stripe (one striping unit from each disk in a RAID group)

Figure 1. Price/performance comparison of different RAID levels

the behavior and advantages of CERC SATA 2s, administrators can compare different types of RAID implementations (see Figure 2).

As Figure 2 shows, CERC SATA 2s can be a cost-effective alternative when the more advanced capabilities of a hardware implementation are not needed. As in a hardware RAID implementation, CERC SATA 2s configures RAID through option ROM.

Role of option ROM in RAID functionality

The role of option ROM is to extend the functionality of the BIOS so that the OS can use vendor-specific hardware or special devices. For CERC SATA 2s, option ROM provides Int 13h support and configures metadata and RAID.

Enabling Int 13h support

Int 13h is a software interrupt that the BIOS provides to the OS to perform disk I/Os. In the Dell CERC SATA 2s implementation, the Int 13h interrupt is not a direct part of the BIOS. Instead, option ROM is loaded at system startup and is invoked to handle disk I/O requests made through Int 13h.

Configuring metadata and RAID

The metadata stores disk configuration information including the type of RAID in use. This data is stored in reserved areas on a disk, which are not available to user data (hardware-based RAID implementations also store metadata in nonvolatile memory). Option ROM is responsible for configuring this metadata on the disks, and system administrators can choose between a RAID-0 and a RAID-1 implementation during system configuration: turn RAID on in the BIOS settings, and then choose the type of RAID using the option ROM interface, which is invoked through the key sequence Ctrl + M. Option ROM then stores the metadata in the reserved areas on the disks, and the drivers read it at system startup.

Historically, hardware vendors have formatted disk metadata differently. The Storage Networking Industry Association (SNIA) Common RAID Disk Data Format (DDF) Technical Working Group was chartered to define a standard data structure describing how data is formatted across the disks in a RAID group. The Common RAID DDF specification is designed to provide standardization and interoperability among different suppliers of RAID technology.[4] The long-term goal is that drives from different systems, vendors, and controllers can be interchanged while preserving the data on those disks.

Key features of the DDF specification include:

- A standard data structure that defines how data should be formatted across the disks in a RAID group
- Interoperability between different suppliers of RAID technology
- A common format for storing RAID configuration information so that data on physical disks can be accessed independent of the RAID controller used
- No requirement for controllers to store this configuration information in the same format in their internal memory

Role of drivers in CERC SATA 2s

Option ROM formats the RAID configuration and writes the metadata onto the SATA drives. After this is accomplished, the drives are ready for OS installation. First, administrators load the drivers for the OS, which provide the ability to read and interpret the metadata previously written by option ROM. The drivers then present the multiple disks as one logical drive to the OS.

RAID implementation	Advantages	Disadvantages	Example(s)
I/O processor (IOP)–based hardware RAID	• Dedicated resources to provide RAID functionality • OS independent • No consumption of host CPU cycles • Large feature set	• Most expensive RAID implementation	Dell PowerEdge Expandable RAID Controller 4, Dual Channel (PERC 4/DC)
I/O controller (IOC)–based hardware RAID	• Easy to embed • Minimal required component space • Less expensive than IOP-based RAID	• Performance capability restricted by small processor and memory bandwidth of controller	Dell PERC 4, imbedded (PERC 4/IM)
OS-based software RAID	• No added cost (included with OS) • Controller independent	• OS dependent • Affected by OS crashes • Not suited for high-CPU-utilization environments because it uses host CPU cycles	Disk Management (Microsoft® Windows® OS), md (Linux® OS)
Driver-based software RAID	• RAID implementation with lowest added cost • Ability to appear to host environment as IOC- or IOP-based RAID	• OS dependent • Not suited for high-CPU-utilization environments because it uses host CPU cycles • Support for RAID-0 and RAID-1 only	CERC SATA 2s

Figure 2. Comparative advantages and disadvantages of RAID implementations

[4] For more information about DDF, see "Solving the RAID Compatibility Puzzle" by Matthew Brisse and Bill Dawkins, Ph.D., in *Dell Power Solutions,* May 2005.

Management of CERC SATA 2s

Management capabilities for CERC SATA 2s depend on the type of Dell server used. CERC SATA 2s is currently available on Dell PowerEdge™ SC420, PowerEdge SC1420, PowerEdge 800, and PowerEdge 1800 servers. On Dell PowerEdge SC420 and PowerEdge SC1420 servers, the RAID Storage Manager (RSM) utility is used. On Dell PowerEdge 800 and PowerEdge 1800 servers, Dell OpenManage™ Storage Services (OMSS) is used.

RSM on Dell PowerEdge SC platforms

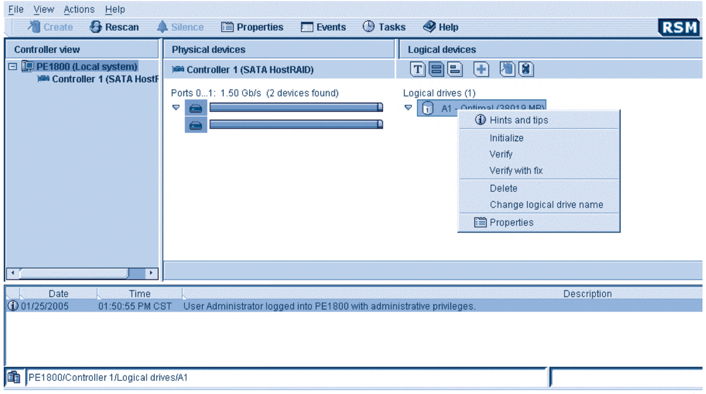
For Dell PowerEdge SC platforms, which are entry-level servers, the RSM utility provides management functionality for the RAID configuration on the Microsoft Windows Server™ 2003 OS (see Figure 3). RSM features include:

- Task management with a scheduler
- Event management
- Multiple graphical user interface (GUI) views
- Documentation with a search facility

Dell OpenManage on Dell PowerEdge platforms

The Dell OpenManage infrastructure provides a comprehensive set of tools, including OMSS and Dell OpenManage Server Administrator (OMSA), to enable efficient systems management. OMSS is accessed through OMSA and can be used to manage the CERC SATA 2s RAID functionality on Dell PowerEdge servers.

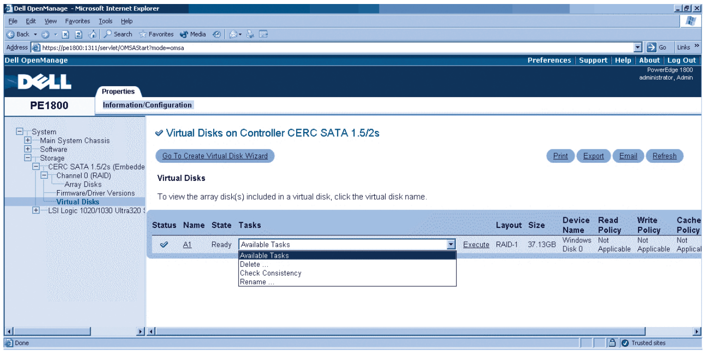
OMSS consists of various components for the management of storage, firmware, and hardware. Using the storage component, administrators can implement controller functions without accessing the BIOS. Controller functions include configuring virtual disks, applying RAID levels, and creating hot spares for data protection. Many controller functions such as rebuilds and troubleshooting can be initiated using OMSS. Most of these functions can be implemented while the server remains online and continues to process requests.



Date	Time	Description
01/25/2005	01:50:55 PM CST	User Administrator logged into PE1800 with administrative privileges.

PE1800/Controller 1/Logical drives/A1

Figure 3. Managing CERC SATA 2s with the RSM utility



Address https://pe1800:1311/servlet/OMSAStart?mode=omsa

PowerEdge 1800
administrator, Admin

PE1800

Virtual Disks on Controller CERC SATA 1.5/2s

Virtual Disks

To view the array disk(s) included in a virtual disk, click the virtual disk name.

Status	Name	State	Tasks		Layout	Size	Device Name	Read Policy	Write Policy	Cache Policy
✓	A1	Ready	Available Tasks	Execute	RAID-1	37.13GB	Windows Disk 0	Not Applicable	Not Applicable	Not Applicab

Figure 4. Managing CERC SATA 2s through the Dell OpenManage Storage Services GUI

The status of storage systems is reported through graphical displays and icons. When a change in status occurs, event traps are sent to the administrator and also maintained in an event log, which can be viewed from the console. In addition, most events generate Simple Network Management Protocol (SNMP) traps, which can be sent to a remote destination.

Figure 4 shows the OMSA/OMSS GUI for managing CERC SATA 2s. In this scenario, a RAID-1 array was created over two SATA disks, and the OS sees this array as a single virtual disk. The OMSS GUI has an Available Tasks drop-down menu, which provides the administrator with various options (delete, check consistency, rename, and so on) for managing the virtual disk.

Cost-effective RAID for low-cost servers

CERC SATA 2s is a low-cost RAID implementation that can provide effective RAID functionality on two SATA drives. It can be used for RAID-0 or RAID-1 and includes management tools for supported Dell PowerEdge and PowerEdge SC platforms. In this way, CERC SATA 2s can provide low-cost servers with basic RAID capabilities while helping cost-conscious organizations run an efficient data center.

Harish Jayakumar is a test engineer in the Dell OpenManage software development and testing organization. He has a master's degree in Computer Science from Arizona State University and a bachelor's degree in Computer Science from the University of Madras, India. His interests include software testing, networking, and operating systems.

Ed Matthews is a senior manager working in the Dell OpenManage software development and testing organization. He has been in the software/IT industry for nearly 30 years, and has prior experience working for semiconductor, telecommunications, storage, and computer manufacturers. He has a B.S. in Physics with Electronics from Liverpool University, England. He is a Chartered Information Technology Professional (CITP) and a Member of the British Computer Society.

SAN Connectivity Kit

for Dell/EMC AX100 and EMC CLARiiON AX100

Switches, HBAs, and software all in one box. **Just add storage!**

Everything you need to build and manage a small SAN … all in one SKU. EMC lab tested for the AX100 storage platform.

QLogic SAN Connectivity Kit™ for AX100

(Order#SAN-C3050-E)

- 1 SANbox 3050-E Fibre Channel Switch
- 2 SANblade QLA200-E Host Bus Adapters (HBAs)
- SANsurfer® Management Software
- 8 SFPs

All for a Great Price!

Add-on Switches and HBAs

SANbox® 3050 Fibre Channel Switch for AX100

(Order#SB3050-08A-E)

- Wizard-based installation takes just minutes
- 8 – 2Gb ports
- SANsurfer® Management Software... Included

SANblade™ QLA200 HBA for AX100

(Order#QLA200-E-SP)

- Wizard-based installation takes just minutes
- Multi-OS support
- SANsurfer® Management Software... Included

For more information, visit **www.Dell.com** and search for **A0434012**.

SIMPLE | **LOW COST** | **SANS**

POWERED BY QLOGIC™

SN0030510-0 Rev A 12/04

Applying a Modular, Layered Approach to

Troubleshooting Tape Drive Hardware

In storage environments, tape drives and libraries can be cost-effective backup and restore devices, but they may incur support costs for troubleshooting and maintenance. This article introduces a modular, layered approach that can ease system administration of tape devices by methodically troubleshooting problems to determine root causes and resolve issues.

BY GAJANAN MUDALIAR

Related Categories:

Dell PowerVault storage

Linux

Microsoft Windows

Novell NetWare

Storage

Tape library

Troubleshooting

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

In the typical enterprise data center, tape drives, tape libraries, and other tape automation devices can be cost-effective workhorses for offline data storage. However, the total cost of ownership for the many tape devices attached to a data center's servers can escalate when support and service costs are factored into the equation. Given these considerations, administrators must take a proactive approach to troubleshooting, determining root causes, and correcting problems.

A basic system for troubleshooting tape drive devices can be categorized in a four-layer stack, similar to a TCP/IP stack. As shown in Figure 1, the predominant layers include the hardware layer, the BIOS layer, the OS layer, and the application layer.

Although the hardware layer and the BIOS layer can be combined into a single layer, various troubleshooting techniques can be carried out independently on each of these layers. This article provides details and troubleshooting techniques for the four layers.

Troubleshooting at the hardware layer

The scope of the hardware layer extends from the tape drive to the SCSI cable connectivity of a SCSI controller. Loose cable connections can cause problems at this layer. Administrators can perform the following troubleshooting procedures to diagnose tape device problems at the hardware layer:

- Check for improper SCSI termination, which can lead to SCSI errors during backups or to the tape

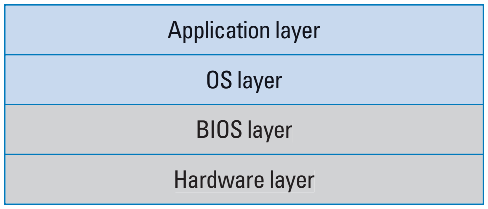


Figure 1. Layered model for modular approach to troubleshooting

device not being detected on the BIOS level.

- Check that the SCSI installation is properly configured—terminators must exist at both ends of the bus. Most host bus adapters (HBAs) are terminated by default, so be sure that the last device in the chain is also terminated.[1]
- Determine how the SCSI device is terminated. Some SCSI devices are internally terminated, which can be disabled through a jumper or a dual in-line package (DIP) switch, while others use an external terminator plug.

Tape drives and libraries usually perform a power-on self-test (POST) operation, in which the system initializes itself and initiates an internal self-test to check for hardware issues. Some tape drives and tape libraries have a display panel, which provides a hex code that corresponds to the hardware error that occurred within the system. The hex code may indicate a self-diagnostic mode or a particular display of blinking or colored lights. Consult the device documentation for more information.

Troubleshooting at the BIOS layer

If a tape drive is not detected on the BIOS level of a controller, the problem usually resides in the controller or the tape drive. Administrators can perform the following troubleshooting procedures to determine why a tape drive is not being detected:

- Change the SCSI ID of the tape drive or library to determine whether the SCSI ID is causing the problem. For example, the tape drive's SCSI ID could conflict with a device on the same SCSI bus, which usually occurs when the SCSI ID of the tape drive is 7—the default SCSI ID of the HBA (SCSI controller). When attaching a tape drive with a SCSI ID above 7 on a narrow SCSI channel, change the SCSI ID to a value below 7.
- Check all the other devices in the SCSI chain. If one of the devices on the same SCSI bus is not responding, it should be separated from the SCSI bus.
- Check the compatibility matrix of the tape drive to determine whether the controller is supported. Commonly used HBAs in Dell™ servers include the Adaptec SCSI Card 39160 and SCSI Card 29160 and the LSI Logic LSI53C1030 embedded SCSI controller.
- Check for incompatibility between standards. For example, if a tape drive is connected to a narrow SCSI channel controller, the narrow SCSI channel can support only seven SCSI IDs.
- Check the configuration of the SCSI host adapter and ensure that the SCSI Disconnection and Parity settings are enabled. Synchronous Negotiation should be disabled. The SCSI transfer rate should be set to 5 Mbps.[2] If hard drives are being duplexed, both SCSI cards should be configured identically.
- Make sure that the SCSI card is set to use Edge Triggering and not Level Triggering. Level Triggering involves the use of shared interrupts, which can cause problems in some systems.

Troubleshooting at the OS layer

If the tape drive is detected on the BIOS level, it should also be detected on the OS level. To facilitate device detection, administrators should proactively check with hardware vendors for updated device drivers.

A basic system for troubleshooting tape drive devices can be categorized in a four-layer stack. The predominant layers include the hardware layer, the BIOS layer, the OS layer, and the application layer.

Microsoft Windows OS

The Microsoft® Management Console (MMC) provides a user interface through which administrators can coordinate systems management for Microsoft Windows® environments. In the Computer Management > Device Manager menu, administrators can view all hardware components within the Windows-based environment. From this menu, they can check whether a tape device is listed. The "?" icon beside a tape device indicates that the device driver is not loaded for that particular device. Device drivers usually can be downloaded from the vendor's Web site or from the CD provided by the vendor.

Novell NetWare OS

System administrators managing systems within Novell® NetWare® environments can troubleshoot tape hardware from the command-line interface. To check whether a tape drive is listed in NetWare, administrators should issue the following command:

```
list storage adapters
```

Sample output of the preceding command is shown in Figure 2.

[1] If both internal and external devices are attached to one host adapter, it may be necessary to disable the termination on the adapter card if it is physically located in the middle of the bus.

[2] Different host adapters will set these parameters in different ways. Some utilize jumpers or DIP switches, while others use a software-based configuration utility. Not all host adapters will be capable of adjusting all of these parameters.

The following command will list only the devices attached to the system:

```
list devices
```

Sample output of the preceding command is shown in Figure 3.

If the device cannot be viewed on the OS level, administrators should issue the following command to scan all the SCSI buses:

```
scan for new devices
```

```
NW2600:list devices
0x0007: [V025-A0-D0:0] SAMSUNG CD-ROM SN-124 N104 [CD]
0x0008: [V024-A5-D1:0] Legacy Floppy [FLOPPY]
0x0009: [V321-A4-D1:0] IBM      ULTRIUM-TD2      37RH [TAPE]
0x000B: [V358-A1-D0:0] FUJITSU MAP3367NC rev:5608 [HDD]
0x000C: [V358-A1-D1:0] FUJITSU MAP3367NC rev:5608 [HDD]
0x000D: [V358-A1-D2:0] FUJITSU MAP3367NC rev:5608 [HDD]
0x000E: [V358-A1-D6:0] Unbound Device Object
```

Figure 3. Sample output for Novell NetWare command to list devices attached to the system

In the device list output, the NetWare OS may not list the tape drive with its actual name or it may list the tape drive as an "unbound device object." In such cases, a proper compatible driver should be loaded so that the output will reflect the inquiry name of the tape drive. The common device driver for all tape devices is nwtape.cdm. For Digital Linear Tape (DLT) tape devices, the common driver is dlttape.cdm.

If the SCSI card is enabled for logical storage units (LUNs), the changer and tape drive may be using the same SCSI ID but different LUNs. If so, administrators must enable LUN support. For example, if using an Adaptec SCSI card, the administrator would add `LUN_ENABLE=FF` to the line that is loading the driver for this card in startup.ncf:

```
LOAD AHA2940.HAM SLOT=HBA slot number LUN_ENABLE=FF
LOAD SYM8XXNW.HAM SLOT=HBA slot number /LUN
```

```
NW2600:list storage adapters
   0x00 [V100-A100] USB UHCI Controller [slot 0]
   0x01 [V025-A0] Novell ATA/IDE Host Adapter Module [slot 10008]
      0x07 [V025-A0-D0:0] SAMSUNG CD-ROM SN-124 N104
   0x02 [V358-A1] LSI_53C1030:10023 [slot 10023]
      0x0E [V358-A1-D6:0] Unbound Device Object
      0x0B [V358-A1-D0:0] FUJITSU MAP3367NC rev:5608
      0x0C [V358-A1-D1:0] FUJITSU MAP3367NC rev:5608
      0x0D [V358-A1-D2:0] FUJITSU MAP3367NC rev:5608
   0x03 [V358-A2] LSI_53C1030:10024 [slot 10024]
   0x04 [V321-A3] Adaptec SCSI Card 39160/3960D - Ultra160 SCSI [slot 201]
   0x05 [V321-A4] Adaptec SCSI Card 39160/3960D - Ultra160 SCSI [slot 202]
      0x0A [V321-A4-D4:0] DELL     PV-132T          227D
      0x09 [V321-A4-D1:0] IBM      ULTRIUM-TD2      37RH
   0x06 [V024-A5] Legacy FLOPPY Controller [slot 0]
      0x08 [V024-A5-D1:0] Legacy Floppy
```

Figure 2. Sample output for Novell NetWare command to list storage devices

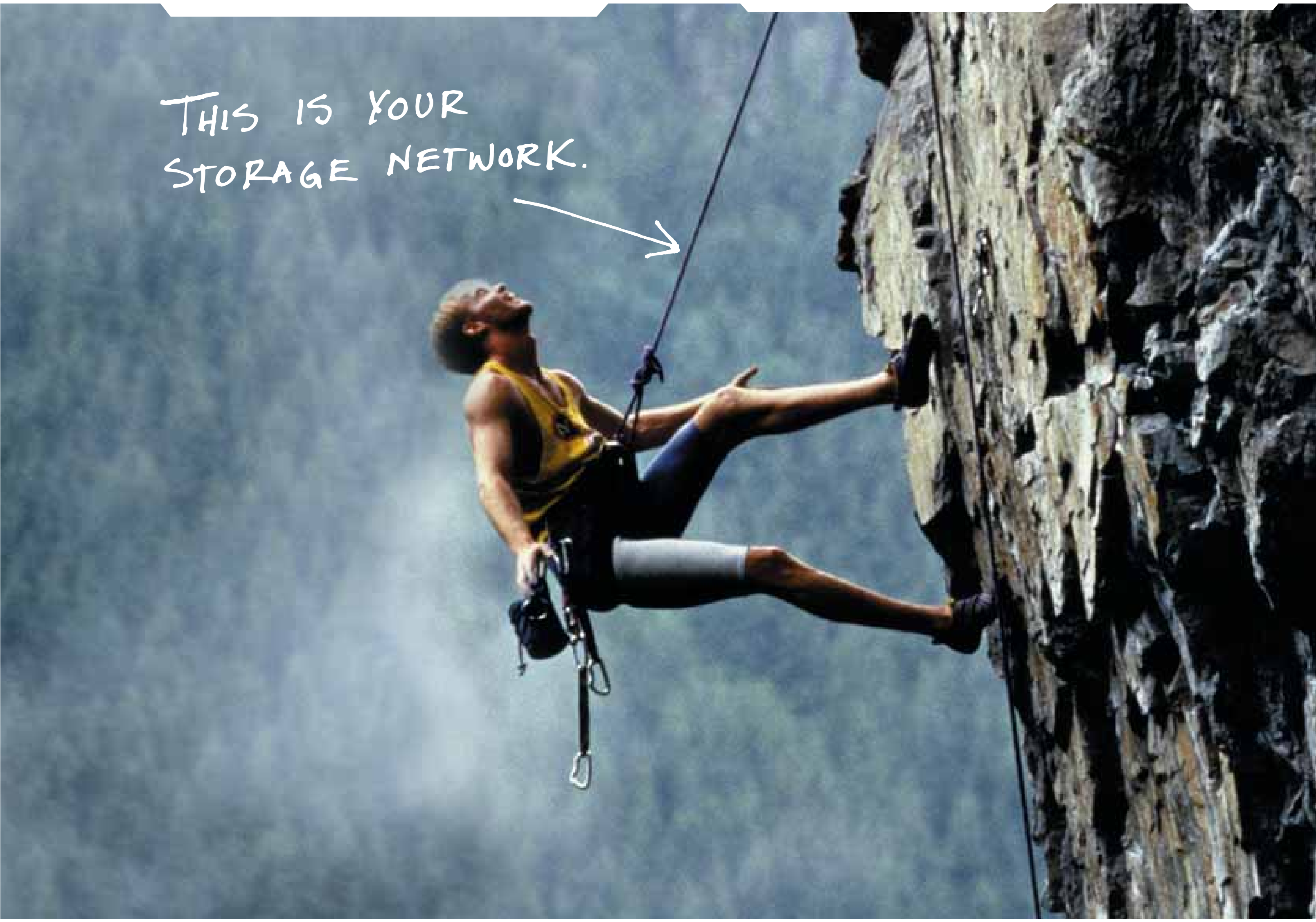
Linux OS

In Linux, /proc/scsi/scsi is the file that shows the inquiry string for the tape drive. The /var/log/dmesg file will have an entry for the tape drive after the kudzu utility is executed and the device driver of the controller is loaded. If a tape library has a changer module and the tape drive and changer have the same SCSI ID but reside in different LUNs—which can occur on Dell PowerVault™ 122T autoloaders with DLT VS 80 technology—the Linux kernel will not update the /proc/scsi/scsi file with the inquiry string. In this case, system administrators should update the file with the following command:

```
echo "scsi add-single-device 2 0 0 13" >
   /proc/scsi/scsi
```

In the preceding command, the four numbers refer to the HBA ID (2), the SCSI bus (0), the SCSI ID (0), and the LUN (13). Devices in Linux are referenced by a device file. A device file can be a raw device file or a logical device file. Because a tape drive is a sequential device, it can be referenced using a raw device path.

The Dell support Web site (support.dell.com) provides numerous tools to help troubleshoot hardware-related tape device issues. Administrators also can initiate a self-diagnostic test to obtain an error code and messages on the front panel of the tape device. The user guide of the tape device can then provide additional information about the error code and troubleshooting steps to help resolve the issue.



Will yours be there when you need it?

Keeping mission-critical data and applications available is of vital importance. And for companies of all sizes, there's no better lifeline than McDATA® multi-capable storage network solutions™. That's because these powerful solutions combine industry-leading hardware, software and services to deliver the scalability, reliability and investment protection that organizations like yours depend on. Just ask more than 80 percent of Fortune 100 companies that rely on McDATA to network the world's business data™.

Learn how you can benefit today from a storage services infrastructure engineered to make the on-demand computing environment a reality. To get your FREE "Business Advantages of a Real-time Storage Services Infrastructure" white paper, visit **www.mcdata.com** today.

McDATA™
Networking the world's business data™

Troubleshooting at the application layer

From the application layer, it is often difficult to drill down to the root cause of problems because most applications do not interact with the hardware directly. Application software generally communicates with the device drivers through the OS kernel, and the device drivers communicate with the firmware of the storage unit. Some applications may have native device drivers loaded to communicate with storage units. Administrators can perform the following troubleshooting procedures to diagnose tape device problems at the application layer:

- Check for wear and tear of the Linear Tape-Open (LTO) tape drive, which can prevent tape media from being populated to maximum capacity.
- Clean the tape head using a cleaning cartridge if a significant amount of debris has accumulated on the head. Media debris can prevent data from being written to a tape cartridge.
- Check whether the device is supported by the application and whether the appropriate device drivers have been installed if the tape hardware is not detected on the application level but is detected on the OS level.
- Check the vendor's Web site to determine whether an updated version of the tape device firmware is available if SCSI errors occur while backing up large amounts of data.[3]

Software vendors typically provide diagnostic utilities to aid in hardware troubleshooting. For example, VERITAS NetBackup software includes a utility through which administrators can initiate a self-diagnostic confidence test on the tape device.

Troubleshooting other general operational issues

Tape drives can break down during a backup or restore procedure. In Windows, administrators would detect this error in a system event log; in Linux and NetWare, the error would appear on the system console or logger screen.

In some cases, to determine the root cause of a failure, administrators must take a SCSI trace and then analyze it on a SCSI analyzer. SCSI analyzers show time diagrams, command listings, and state listings. When using SCSI analyzers, administrators should be familiar with all the respective SCSI commands so that they can check the conditions and the trace communications over the SCSI bus.

The Microsoft Management Console provides a user interface through which administrators can coordinate systems management for Microsoft Windows environments.

Whenever SCSI errors or hardware errors occur in the backup device, the target—the device that receives service requests for processing—sets a unit attention (UA) flag, which eventually follows a SCSI reset condition. Administrators can determine the cause of a SCSI bus reset by taking a SCSI trace of the communication path between the target and the initiator—a SCSI device containing application clients and SCSI ports that originate device service and task management requests to be processed by a SCSI target device. Obtaining the sense data information from the SCSI trace session can show which flags have been set and help administrators identify a clear path to problem resolution.

System administrators managing systems within Novell NetWare environments can troubleshoot tape hardware from the command-line interface.

Efficiently resolving tape hardware problems

Administrators can enhance the process of troubleshooting tape devices by following the modular, layered approach described in this article. To determine root causes rapidly, administrators should first categorize the condition as a hardware issue or a software issue, then determine the layer at which to begin troubleshooting. This method can help administrators find an efficient path to problem resolution.

Gajanan Mudaliar is an engineering analyst on the tape storage engineering team in the Dell Product Group. His primary area of expertise is testing and troubleshooting tape hardware and automation devices. Prior to joining Dell, Gajanan had seven years of IT experience in technical support and system administration for enterprise storage installations. He has a B.E. in Electronics and Telecommunication from the University of Kolhapur in India.

FOR MORE INFORMATION

Dell support:
support.dell.com

VERITAS support:
support.veritas.com

[3] As a general safety rule, do not disconnect or power down the tape device during the firmware upgrade.

Automating Microsoft SQL Server 2000 Online Backup with the EMC SnapView Integration Module

The Microsoft® SQL Server 2000 Virtual Device Interface is designed to help organizations achieve fast backups and restores of SQL Server databases without affecting production workloads, thus simplifying system administration and backup management. This article describes how EMC® SnapView™ Integration Module for SQL Server can help automate the online backup process for Microsoft SQL Server 2000 databases.

BY ANANDA SANKARAN AND KEVIN GUINN

Related Categories:

Database

Dell PowerEdge servers

Dell/EMC storage

Disaster recovery

Fibre Channel

McDATA

Microsoft SQL Server

Microsoft Windows

Storage software

Storage

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

A common challenge in creating an effective backup strategy is minimizing the window required for backup and restore operations on database volumes. Common solutions to this challenge include using point-in-time copies, or *snapshots*, as source volumes for backup operations. Business continuance volumes (BCVs)—or *clones*—are an alternative type of snapshot that can create a full duplicate of the data.

EMC SnapView software is designed to create snapshots and clones quickly and nondisruptively in database environments. The EMC SnapView Integration Module for SQL Server (SIMS) helps automate the process of stabilizing the database volume on the production host and making an image of the quiescent data available to the backup host. It uses the Microsoft SQL Server 2000 Virtual Device Interface (VDI) and manages SnapView clones and snapshots on Dell/EMC CX series storage arrays. Advantages of using SIMS for SQL Server backup include high database availability, excellent database backup performance, and fast database recovery times.

SQL Server Virtual Device Interface

The SQL Server 2000 VDI is a high-performance backup interface. It enables third-party backup solutions to interoperate with SQL Server to perform high-speed online data backups. The VDI can help achieve high backup throughput and fast restore times with minimal overhead on a production workload.

The VDI provides support for two types of backup: conventional online backups and snapshot backups. Conventional online backups can use the full range of SQL Server backup and restore features. Snapshot backups are performed using the underlying storage vendor's snapshot technology—either split-mirror or copy-on-write—and are limited to full database and file (or file group) backups only. Snapshot backups can be combined with database differential, file differential, and transaction log backups to recover a database to a specific point in time. The advantages of using hardware-based snapshots for backup are extremely fast backup and restore times with no impact on production workload, which helps organizations achieve highly available database systems.

Third-party backup software usually includes a snapshot provider that interacts with SQL Server through the VDI. A snapshot provider creates a near-instant copy of the database files residing on the volume being captured. As the snapshot provider interacts with SQL Server through the VDI, the backup software first captures the backup data set's metadata and then performs the snapshot. When performing a backup, the backup software issues a backup command to the SQL Server VDI application programming interface (API). The backup software receives the backup data set's metadata from SQL Server through the VDI but does not receive the backup data set itself (the data and log files), because the snapshot of the database volume already contains the actual data and log files. The backup data set's metadata is required to restore the database from the snapshot.

Snapshot duration is the time elapsed between SQL Server issuing a snapshot command to the snapshot provider and the provider returning a successful completion. Writes to the database files being captured in the snapshot are suspended for the duration of the snapshot—and thus the snapshot provider must complete the snapshot operations quickly to avoid any impact on the SQL Server production workload. The snapshot provider must indicate completion as soon as SQL Server writes to the database files become possible, while ensuring that the snapshot remains consistent.

The VDI provides an interface to SQL Server that allows a backup solution to act as a virtual backup device (see Figure 1). SQL Server writes to the virtual device during backup operations, and reads from the virtual device during restore operations. During a backup job, the snapshot provider backs up the database by accepting a VDI data stream; during a restore job, SQL Server receives the VDI stream from the snapshot provider to recover the database. The VDI enables high-speed data transfer between SQL Server and the snapshot provider with minimal overhead.

Split-mirror versus copy-on-write snapshot technology

Generally, vendors offer two types of technology to produce snapshots using the VDI: split-mirror and copy-on-write.

- **Split-mirror snapshots:** A split-mirror snapshot is based on storage mirroring solutions and is an exact copy of the original database volume. This kind of snapshot is most commonly referred to as a BCV or clone. In the split-mirror approach, one image is frozen and no longer updated, while the other continues to be live and therefore continues to be updated. The snapshot obtained after splitting the mirror and the metadata (which is created by SQL Server to describe the snapshot) constitute the database backup. Split-mirror snapshots reduce the time needed for the restoration of a full database backup but add an extra cost to the backup solution because additional mirror disks are required. Both the EMC SnapView package and the EMC SnapView clone provider are required to create BCVs on Dell/EMC CX series storage arrays.
- **Copy-on-write snapshots:** A copy-on-write snapshot is a copy of all the original disk blocks in a database that have changed since the image was created. The snapshot and the production database share the majority of their disk blocks and differ only in those disk blocks that have been modified since the snapshot was created. Copy-on-write snapshots are not as fast as split-mirror snapshots but are less expensive because no special hardware is required—and they incur minimal disk space as overhead for maintaining the snapshot. The EMC SnapView package is required to create these snapshots on Dell/EMC CX series storage arrays.

EMC SnapView software is designed to create snapshots and clones quickly and nondisruptively in database environments.

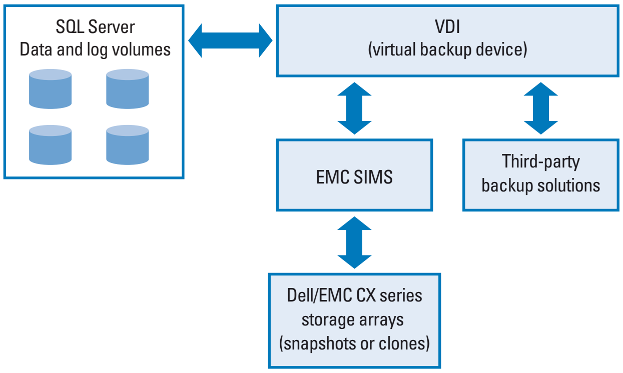


Figure 1. SQL Server, the VDI, and backup solutions

SnapView Integration Module for SQL Server

SQL Server databases residing on Dell/EMC CX series storage arrays can be backed up online using the EMC SIMS module. SIMS is designed to create snapshots and clones of the SQL Server database by using EMC SnapView technology and the Microsoft SQL Server 2000 VDI. The database copies can then be backed up to disks or appropriate tape solutions.

With EMC SIMS, the process of creating the database copy and backing it up occurs while the server is online. Normal backup solutions incur a large backup window by placing the SQL Server in backup mode, potentially for several hours depending on the size of the files that are backed up. With SIMS, SQL Server is placed in backup mode only for the few seconds required by the SnapView process to complete the snapshot or clone operations. This approach requires only a small backup window and helps keep SQL Server primarily in its normal operating mode. In addition, the overhead incurred on the SQL Server production host is minimal because it is not involved in moving the data for backup. Instead, this task

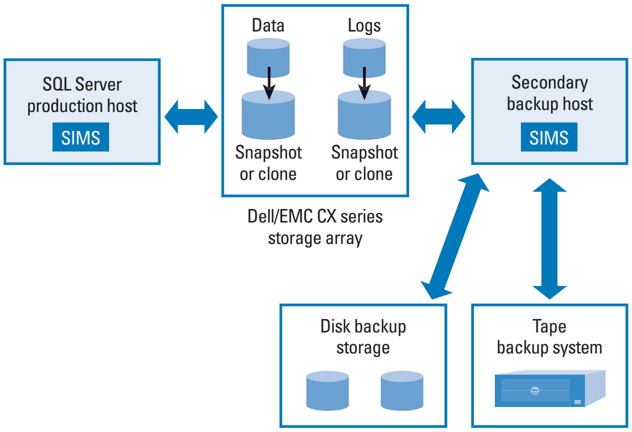


Figure 2. SIMS-based backup

is offloaded to the CX series storage array, and the production host can operate normally during backup.

SIMS provides the following three backup job types:

- **Snapshot volume job type:** For the snapshot volume (SV) job type, SIMS creates SnapView snapshots for the backup process. SIMS creates and activates snapshots of the SQL Server database source logical storage units (LUNs) mounted on the production host. It provides options for backing up the snapshots to disk or to a third-party tape backup solution using a secondary backup host (see Figure 2). This job type uses minimal disk space because it backs up only the minimal changes to the data that occurred after the initial creation of the point-in-time copy. During backup, SIMS checks SQL Server, starts the VDI process, and flushes the production host file system. The VDI process freezes the database for writes and creates metadata for the backup. SIMS then completes the snapshot process on the storage array and exits the VDI process. The snapshots that contain the stable image of the database data are accessed by the secondary backup host and can be backed up to disk or passed to a third-party tape backup solution.

With EMC SIMS, the process of creating the database copy and backing it up occurs while the server is online.

- **BCV job type:** For the BCV job type, SIMS uses preexisting SnapView clones for the backup process. SIMS synchronizes clones of the database source LUNs mounted on the production host. The clones provide a complete copy of the database source LUNs and can be used for rapid recovery of the database on the secondary backup host. During backup, SIMS checks SQL Server, synchronizes the clones, starts the VDI process, and performs a check on the database to flush the database files. SIMS then administratively fractures (that is, splits the mirror of) the clone LUNs on the storage array and exits the VDI process. The clones that contain the database data are then accessed on the secondary backup host and backed up to disk or passed to a third-party tape backup solution.
- **BCV-SV job type:** This hybrid job type is similar to the BCV job type except that SIMS starts and activates snapshots of the cloned LUNs. The clones can be used for rapid recovery, and the snapshots can be backed up to disk or passed to a third-party tape backup system. This job type provides both a complete copy and a point-in-time copy of the database data.

SIMS requirements

SIMS interacts with several other EMC software components on the host and the storage array. These interactions allow SIMS to manage snapshots and clones on the Dell/EMC CX series storage array, quiesce SQL Server 2000 database volumes on the production host, and present a stable image of the SQL Server volumes to a backup host.

Figures 3 through 5 indicate which software components were tested for this article on the SQL Server production host, the secondary

Component	Version
Microsoft SQL Server 2000	Service Pack 3
EMC PowerPath® software (recommended)	3.0.6
EMC Navisphere® Host Agent	6.6.0.3.8
EMC SIMS	1.00
EMC SnapView admsnap utility	2.3.0.0.9
EMC Navisphere command-line interface (CLI)	6.6.0.3.8
Java 2 Runtime Environment (J2RE)	1.4.2_04

Figure 3. SQL Server production host software requirements

Component	Version
EMC PowerPath software	3.0.6
EMC Navisphere Host Agent	6.6.0.3.8
EMC SIMS	1.00
EMC SnapView admsnap utility	2.3.0.0.9
EMC Navisphere CLI	6.6.0.3.8
J2RE	1.4.2_04
Tape backup software (recommended)	Any supported tape backup application

Figure 4. Secondary backup host software requirements

Component	Version
EMC FLARE™ operating environment	02.06.x00.5.01x
EMC Navisphere software	6.6.0.3.18
EMC Access Logix™ software	01.01.5.001
EMC SnapView software	02.03.5.004
EMC SnapView clone provider	6.6.0.3.6
EMC Navisphere CLI provider	6.6.0.3.1
EMC Navisphere management server	6.6.0.5.0
EMC Navisphere Manager GUI (recommended)	6.6.0.3.6
EMC SnapView GUI (recommended)	6.6.0.4.1

Figure 5. Dell/EMC CX series storage array software requirements

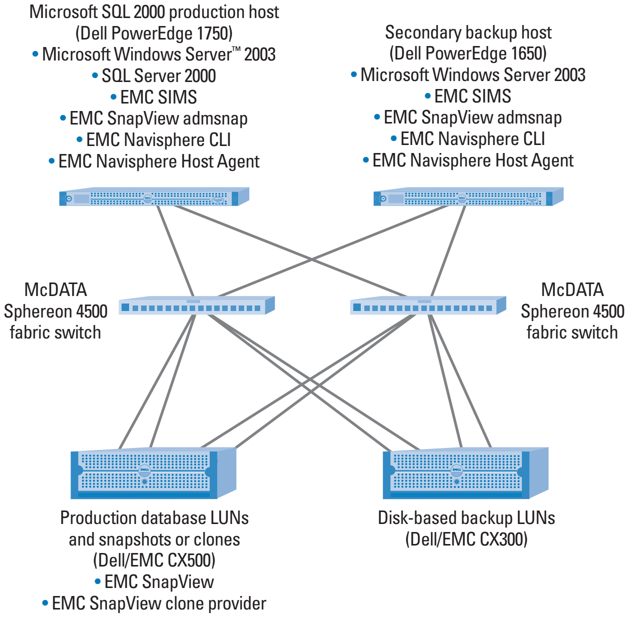
backup host, and the Dell/EMC CX series storage array in May 2004. The indicated versions are representative of the software components that are necessary to implement a backup solution for SQL Server 2000 using EMC SIMS.

Integrated SQL Server 2000 online backup

Figure 6 illustrates an online backup scenario for SQL Server 2000 on Dell/EMC hardware. In this scenario, the backup administrator configures a backup job schedule (either the SV or BCV job type) on the secondary backup host. When SIMS runs the scheduled job, it interacts with SQL Server on the production host through the VDI. SQL Server then goes into online backup mode until SIMS completes the snapshot or clone process (depending on the job type). Next, SIMS performs the snapshot or clone process on the Dell/EMC storage array and communicates the completion of the process to SQL Server through the VDI. Upon receiving notification that the backup process has completed, SQL Server resumes its normal operation. Resuming normal operations takes only a few seconds. The snapshots or clones created for the database data can then be accessed on the secondary backup host. Administrators can back up to disk on a different storage array or to a third-party tape backup solution.

Reliable, simplified database backup systems

IT administrators can build reliable, simplified database backup systems using standards-based enterprise hardware such as Dell™ PowerEdge™ servers, Dell/EMC storage arrays, and Dell PowerVault™ tape libraries. By interacting with the Microsoft SQL Server 2000 VDI and managing SnapView snapshots and clones on Dell/EMC CX series storage arrays, EMC SIMS is designed to perform online backups of SQL Server 2000 production databases. EMC SIMS can facilitate online backups by automating the processes of stabilizing the production data, producing an image of the stabilized data, and presenting that image to a secondary backup host. The backup host can save the data to disk or use a third-party tape software application to save the data to tape. This approach simplifies both system administration and backup management.


Microsoft SQL 2000 production host
(Dell PowerEdge 1750)
• Microsoft Windows Server™ 2003
• SQL Server 2000
• EMC SIMS
• EMC SnapView admsnap
• EMC Navisphere CLI
• EMC Navisphere Host Agent
Secondary backup host
(Dell PowerEdge 1650)
• Microsoft Windows Server 2003
• EMC SIMS
• EMC SnapView admsnap
• EMC Navisphere CLI
• EMC Navisphere Host Agent
McDATA Sphereon 4500 fabric switch
McDATA Sphereon 4500 fabric switch
Production database LUNs and snapshots or clones
(Dell/EMC CX500)
• EMC SnapView
• EMC SnapView clone provider
Disk-based backup LUNs
(Dell/EMC CX300)

Note: The CX500 is accessible by both hosts. The CX300 is accessible only by the secondary backup host. Accessibility is controlled by zoning on the Fibre Channel switches.

Figure 6. Integrated backup configuration using SIMS

Ananda Sankaran is a systems engineer in the High-Availability Cluster Development Group at Dell. His current interests related to high-availability clustering include storage systems, application performance, business continuity, and cluster management. Ananda has a master's degree in Computer Science from Texas A&M University.

Kevin Guinn is a systems engineer in the High-Availability Cluster Development Group at Dell. His current interests include storage management and business continuity. Kevin is a Microsoft Certified Systems Engineer (MCSE) and has a B.S. in Mechanical Engineering from The University of Texas at Austin.

FOR MORE INFORMATION

EMC SnapView:
www.emc.com/products/software/snapview2.jsp

SQL Server:
www.microsoft.com/sql

SQL Server availability:
www.microsoft.com/sql/evaluation/features/maximize.asp
www.microsoft.com/sql/evaluation/features/reliable.asp

Introducing Microsoft Windows Server 2003 x64 Editions for the Intel EM64T Platform

The launch of Microsoft® Windows Server™ 2003 x64 Editions marks a milestone toward the acceptance of 64-bit technology in mainstream computing. This 64-bit platform offers several enhanced features and advantages, including the opportunity for IT environments to investigate 64-bit technology without giving up 32-bit applications. This article provides an overview of Windows Server 2003 x64 Editions and their distinguishing features, and compares them with 32-bit Microsoft Windows® operating systems.

BY RANJITH PURUSH AND CHIP WEBB

Related Categories:

Dell PowerEdge servers

Intel EM64T processors

Microsoft Windows

Microsoft Windows Server 2003 x64 Editions

Operating system (OS)

Planning

System deployment

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

Microsoft Windows Server 2003 x64 Editions—also referred to as x64—are designed for servers based on 64-bit extended architecture. This article addresses key features of the Intel® Extended Memory 64 Technology (EM64T)[1] architecture to help provide a better understanding of the behavior of Windows Server 2003 x64 Editions on this platform. Note that this OS does not support 64-bit Intel Itanium® processors; the Windows Server 2003 for 64-Bit Itanium-based Systems OS is designed for the Itanium platform. References in this article to 64-bit technology refer to EM64T and are not applicable to 64-bit Itanium-based hardware or Windows Server 2003 for 64-Bit Itanium-based Systems.

Compatibility with 32-bit applications

A key feature of the 64-bit extended architecture is the capability to run both 32-bit and 64-bit applications. Dell™ PowerEdge™ servers with Intel EM64T–capable processors[2] support this capability by using the different operating modes and sub-modes within the Intel EM64T architecture. Figure 1 shows the different operating modes of the 64-bit extended processor.

Long mode is the 64-bit mode of the 64-bit extended processor, and this mode supports x64. Legacy mode supports 32-bit x86-based operating systems such as 32-bit Windows Server 2003. When a Windows Server 2003 x64 Editions OS is booting up, the OS automatically[3] switches from Legacy mode, which is the processor's initial mode, to Long mode.[4] A 32-bit x86-based OS such as 32-bit Windows Server 2003 cannot switch the processor mode to Long mode—the processor will always be in Legacy mode for these operating systems.

Long mode has two sub-modes: Compatibility and 64-bit. Compatibility mode allows Windows Server 2003

[1] For more information about the hardware capabilities and features of Intel EM64T, visit www.intel.com/technology/64bitextensions.

[2] For more information about installing Windows Server 2003 x64 Editions on supported Dell PowerEdge servers, see "Deploying Microsoft Windows Server 2003 x64 Editions on Dell PowerEdge Servers" by Ranjith Purush and Sandhya Senapathi in *Dell Power Solutions,* May 2005.

[3] NT Loader (Ntldr) performs the switching function. No special switches are required in the boot.ini file to accomplish this process. Unlike the Windows Server 2003 for 64-Bit Itanium-based Systems OS, Windows Server 2003 x64 Editions do not require a separate partition to store boot code.

[4] Windows Server 2003 x64 Editions enable Long mode of an EM64T-capable processor by setting the Long Mode Active (LMA) control bit, which is bit 10 of the extended feature enable register (IA32_EFER).

x64 Editions to run existing 32-bit applications natively[5] on a 64-bit processor without recompilation. The capability to run 32-bit applications natively on the 64-bit processor is a distinguishing feature of x64, and is possible because the x64 instruction set is an extension of the industry-standard x86 architecture. Applications running in Compatibility mode use 32-bit data operands and 32-bit addressing that allow access only to the first 4 GB of virtual address space. However, as true 64-bit operating systems, Windows Server 2003 x64 Editions are designed to use 64 bits for all work—even while a 32-bit application is using only 32 bits. The 64-bit mode supports 64-bit applications.

Processor operating modes and sub-modes		Required OS	Application recompilation required?	Default address size	Default data size	Register extensions allowed?
Long mode	64-bit mode	x64 OS	Yes	64-bit	32-bit	Yes (64-bit GPR*** width)
	Compatibility mode*		No**	32-bit	32-bit	No (32-bit GPR width)
Legacy mode		32-bit x86-based OS	No	32-bit	32-bit	No (32-bit GPR width)

**The EM64T-capable processors support 16-bit applications in Compatibility mode. However, Windows Server 2003 x64 Editions do not support 16-bit applications.*
***The 32-bit applications that are not recompiled to 64-bit will run on x64 in Compatibility mode.*
****The 32-bit x86 architecture offers only eight 32-bit general-purpose registers (GPRs). Intel EM64T increases the number of GPRs to 16, and all GPRs in EM64T-capable processors are 64-bit.*

Figure 1. Different operating modes of Intel EM64T–capable processors

	Legacy mode	Long mode	
		Compatibility mode	64-bit mode
Applications	32-bit	32-bit	64-bit
Drivers	32-bit	64-bit	64-bit
OS	32-bit	64-bit	64-bit
Flat address space	4 GB	4 GB	16 TB
General-purpose registers	32-bit	32-bit	64-bit

Figure 2. Applications, drivers, and operating systems supported in the different modes of Intel EM64T–capable processors

	32-bit		x64	
Windows OS edition*	Supported physical memory	Supported physical processors	Supported physical memory	Supported physical processors
Windows XP Professional**	4 GB	1–2	128 GB	1–2
Windows Server 2003, Standard Edition	4 GB	1–4	32 GB	1–4
Windows Server 2003, Enterprise Edition	32 GB/ 64 GB***	1–8	1 TB	1–8

**Microsoft offers four x64 editions: Windows XP Professional x64 Edition and the Standard, Enterprise, and Datacenter editions of Windows Server 2003 x64. At press time, Dell did not support Windows Server 2003, Datacenter x64 Edition.*
*** This article focuses on server operating systems and thus does not include Windows XP Professional. For more information about Windows XP Professional x64 Edition, visit www.microsoft.com/windowsxp/64bit/default.mspx.*
**** Memory support increased to 64 GB in 32-bit Windows Server 2003, Enterprise Edition, with Service Pack 1 (SP1).*

Figure 3. Comparison of processor and memory support for 32-bit and x64 Windows operating systems

The determination of whether the processor is in 64-bit mode or Compatibility mode is based on the currently executing code segment.[6] In Windows Server 2003 x64 Editions, a 32-bit application[7] will always be in Compatibility mode and a 64-bit application will always be in 64-bit mode. The x64 architecture and components are discussed in further detail later in this article.

Figure 2 shows the types of applications, drivers, and operating systems that are supported in the different modes of the x64-capable processor. Legacy mode preserves the binary compatibility with 32-bit applications and supports only the Legacy Protected mode, which is equivalent to an x86 32-bit protected-mode environment.

Extended memory support

The 64-bit OS significantly extends physical memory and memory allocation limits. Figures 3 and 4 summarize the memory support in Windows Server 2003 x64 Editions and compare it to 32-bit Windows operating systems. Note that x64 enables increased physical memory support—up to 32 GB for Standard x64 Edition and up to 1 TB for Enterprise x64 Edition.

Virtual memory comprises a large address space used by processes and applications. Some parts of the virtual memory space may be located in physical memory, while other parts of virtual memory may be located on storage media such as hard drives. Windows Server 2003 x64 Editions allow 64-bit processes and applications to use a

Memory limits	32-bit Windows Server	Windows Server 2003 x64 Editions
Total virtual address space	4 GB	16 TB
Virtual address space per 32-bit process	2 GB (3 GB with /3GB switch)	2 GB/4 GB
Virtual address space per 64-bit process	Not applicable	8 TB
Paged pool	470 MB (650 MB in Windows Server 2003 SP1)	128 GB
Non-paged pool	256 MB	128 GB
System cache	1 GB	1 TB

Figure 4. Comparison of memory allocation limits for 32-bit and x64 Windows operating systems

[5] The 64-bit Itanium processors do not run 32-bit applications natively. Limited support for 32-bit applications on Itanium processors is made possible by the Windows Server 2003 for 64-Bit Itanium-based Systems OS—the processor has an x86 emulation mode through which Windows can execute x86 programs, albeit with a substantial performance penalty.

[6] The processor switches modes based on the values of two bits (L and D) in the segment descriptor specified by the selector in the Code Segment (CS) register. The code segment of a 32-bit process is set to have L=0 and D=1, while a 64-bit process will set L=1 and D=0.

[7] More specifically, a process cannot switch from 64-bit mode to Compatibility mode and vice versa. However, a process in one mode can launch a process in the other mode.

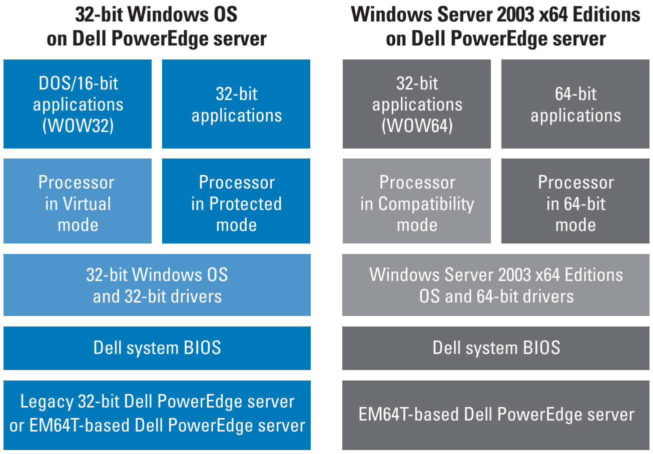

32-bit Windows OS on Dell PowerEdge server
DOS/16-bit applications (WOW32)
32-bit applications
Processor in Virtual mode
Processor in Protected mode
32-bit Windows OS and 32-bit drivers
Dell system BIOS
Legacy 32-bit Dell PowerEdge server or EM64T-based Dell PowerEdge server
Windows Server 2003 x64 Editions on Dell PowerEdge server
32-bit applications (WOW64)
64-bit applications
Processor in Compatibility mode
Processor in 64-bit mode
Windows Server 2003 x64 Editions OS and 64-bit drivers
Dell system BIOS
EM64T-based Dell PowerEdge server

Figure 5. Comparison of requirements for drivers and applications in 32-bit and x64 Windows operating systems

single contiguous memory space and access locations anywhere in the linear 64-bit address range.[8] This capability allows the 64-bit OS to run memory-intensive applications, such as Microsoft SQL Server, more efficiently than the 32-bit OS. The 32-bit processes running on x64 may be allocated up to 2 GB of virtual address space, and administrators can increase this limit to 4 GB by compiling the 32-bit applications with the `/LARGEADDRESSAWARE` switch. Windows Server 2003 x64 Editions do not support the `/3GB` switch.[9]

Driver requirements for x64

Figure 5 summarizes the application and driver support requirements for Windows Server 2003 x64 Editions and shows the difference between an x64 OS and a 32-bit OS. A key distinction is that Windows Server 2003 x64 Editions require all drivers—including all device drivers—to be 64-bit drivers.[10] All 32-bit kernel-mode drivers must be ported to 64-bit.[11] Because Windows Server 2003 x64 Editions are new, Microsoft has transformed certain best-practices recommendations for 32-bit operating systems into mandatory requirements for x64, as follows:

- **x64 device drivers:** All x64 device drivers should have x64-specific decorators in the driver .inf files. For more information, visit www.microsoft.com/whdc/device/storage/F6dirs.mspx and www.microsoft.com/whdc/driver/install/64INF_reqs.mspx.
- **Kernel-mode drivers:** Kernel-mode drivers cannot patch (extend or replace) kernel services by modifying system service tables, modifying the Interrupt Descriptor Table (IDT) and Global Descriptor Table (GDT), and so forth. For more information on this requirement, visit www.microsoft.com/whdc/driver/kernel/64bitpatching.mspx.

The WOW64 subsystem: The replacement for WOW32

Another important distinction in Windows Server 2003 x64 Editions is the removal of the Virtual mode (Windows On Windows 32, or WOW32) that allows 16-bit application support on 32-bit systems. Because of this change, 16-bit applications or applications that have 16-bit components are not supported on x64. Attempting to run a 16-bit application on an x64 OS will generate an informational message that informs the user of the incompatibility. Because many mainstream 32-bit applications have 16-bit installer components, Microsoft has provided a work-around that allows x64 to support such 32-bit applications.[12] Windows Server 2003 x64 Editions recognize these supported 16-bit programs and automatically substitute them with their 32-bit versions. Windows Server 2003 x64 Editions support the use of 32-bit Windows installers for 32-bit applications; however, 64-bit applications must have 64-bit installers to properly access the 64-bit folders and native 64-bit registry hives.

The WOW32 subsystem has been replaced with the WOW64 (Windows On Windows 64) subsystem in x64 (see Figure 5). The WOW64 subsystem, which is a component in Windows Server 2003 x64 Editions, is the 32-bit Windows emulation layer that ensures binary compatibility with 32-bit Windows applications and allows 32-bit x86-based applications to run on 64-bit Windows Server 2003 x64 Editions.

Features of the WOW64 subsystem such as the file system redirector and the registry redirector allow 32-bit applications to run on an x64 system with complete transparency.

The core of WOW64 consists of the following three dynamic-link libraries (DLLs):

- **Wow64.dll:** Manages process and thread initialization as well as exception dispatching to 32-bit code; intercepts base system calls (exported by Ntoskrnl.exe); and implements file system and registry redirection (discussed later in this section) and registry reflection

[8] Current Windows Server 2003 x64 Editions implementations use 46-bit virtual addresses. Future implementations are expected to increase this limit as PC hardware evolves.

[9] The 32-bit Windows operating systems allow the use of the `/3GB` switch in the boot.ini file to increase the virtual address space for a 32-bit process from 2 GB to 3 GB. Because this effectively restricts the address space available to the kernel to 1 GB, there may be a negative performance impact from using the `/3GB` switch. Only applications compiled with the `/LARGEADDRESSAWARE` switch can use more than 2 GB of virtual address space.

[10] Note that the binary versions of 64-bit drivers and 64-bit applications for Windows Server 2003 x64 Editions are very different from those for the Windows Server 2003 for 64-Bit Itanium-based Systems OS.

[11] For more information about porting drivers to 64-bit, visit msdn.microsoft.com/library/default.asp?url=/library/en-us/kmarch/hh/kmarch/Other_394c38ae-a3e6-45fb-87f2-c3e227cb6b7c.xml.asp.

[12] A list of supported 16-bit installer programs is available in the registry at the following location on an x64 system: HKLM\Software\Microsoft\Windows NT\CurrentVersion\NtVdm64.

Reprinted from *Dell Power Solutions,* May 2005. Copyright © 2005 Dell Inc. All rights reserved.

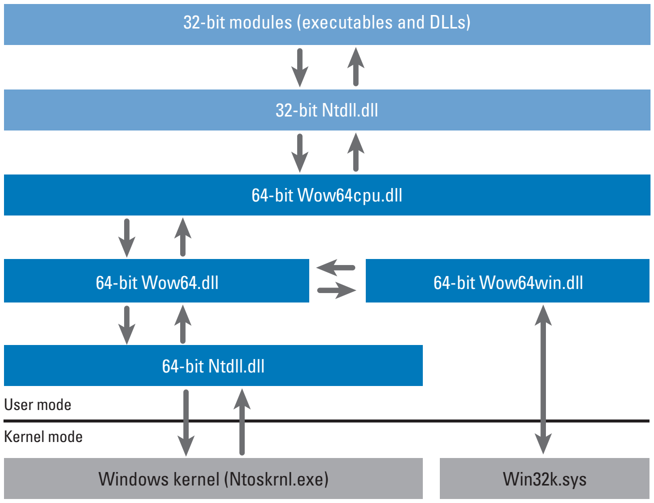


Figure 6. Interaction between 32-bit applications and the 64-bit kernel via WOW64

- **Wow64win.dll:** Intercepts graphical user interface (GUI) system calls (exported by Win32k.sys)
- **Wow64cpu.dll:** Manages the 32-bit CPU context of each running thread inside WOW64 and provides processor architecture–specific support for switching CPU modes from 32-bit to 64-bit and vice versa

Figure 6 shows the interaction between 32-bit applications and the 64-bit kernel via the WOW64 layer. The WOW64 subsystem intercepts system calls[13] from the 32-bit application, transitions from Compatibility mode to 64-bit mode, converts all 32-bit data structures into 64-bit aligned data structures, issues the native 64-bit system call, writes back any output data from the 64-bit system call, and returns back to 32-bit Compatibility mode. This process of conversion performed by the WOW64 layer is commonly referred to as *thunking* and is transparent to the software developer in most cases.

Restrictions and limitations of the WOW64 layer include the following:

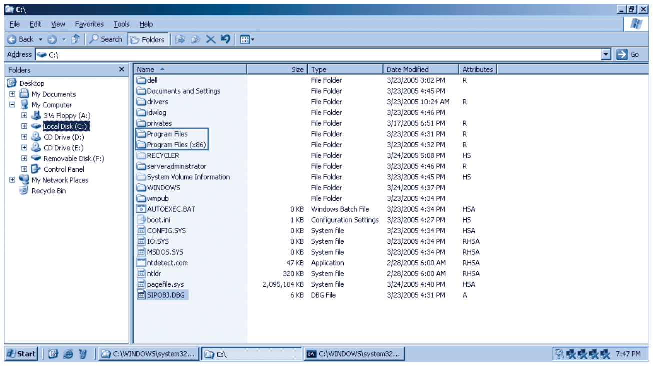
- WOW64 processes (32-bit processes) cannot load 64-bit DLLs (except for the core 64-bit Ntdll.dll and the 64-bit WOW64 binaries—Wow64.dll, Wow64win.dll, and Wow64cpu.dll); native 64-bit processes cannot load 32-bit DLLs.
- WOW64 does not support 16-bit applications, including the 16-bit WINNT.exe application.
- WOW64 cannot handle input buffer for the DeviceIoControl function.

The capability of Windows Server 2003 x64 Editions to support both 32-bit and 64-bit applications also introduces complexities in the coexistence of 32-bit and 64-bit versions of drivers, system files, and registry values. The x64 platform is designed to prevent issues relating to the coexistence of 32-bit and 64-bit applications through the implementation of the file system redirector, registry redirector, and registry reflection features.

Windows Server 2003 x64 Editions implement two Program Files directories, as shown in Figure 7. The Program Files folder hosts the 64-bit application program files and the Program Files (x86) folder hosts the 32-bit versions of the program files. These directories enable 32-bit and 64-bit applications to coexist without conflict on an x64 system. There are also two System folders, as shown in Figure 8. The System32 folder hosts the 64-bit system files, and the SysWOW64 folder contains the 32-bit files. The WOW64 subsystem can see only the Program Files (x86) and SysWOW64 folders.

File system redirector and registry redirector

Features of the WOW64 subsystem such as the file system redirector and the registry redirector allow 32-bit applications to run on an x64 system with complete transparency. Any 32-bit application that attempts to access the Program Files or System32 folder will automatically be resolved to the Program Files (x86) or SysWOW64 folder, respectively, by the WOW64 subsystem. This automatic path resolution, which depends on whether the application is 32-bit or 64-bit, is implemented by WOW64 and is known as file system redirection.[14] To facilitate backward compatibility, the following

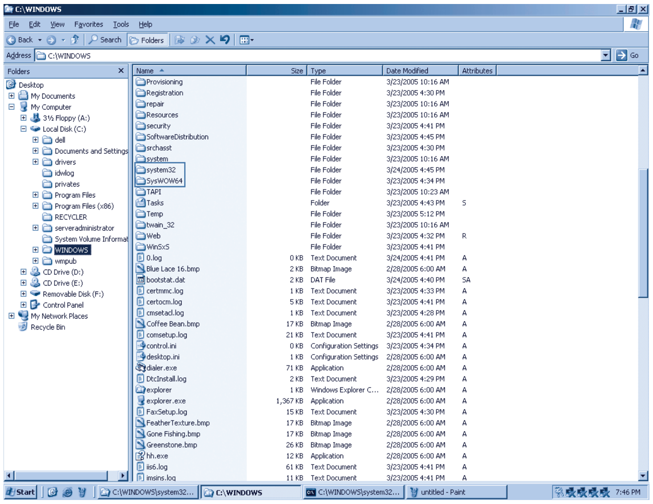


Name	Size	Type	Date Modified	Attributes
dell		File Folder	3/23/2005 3:02 PM	R
Documents and Settings		File Folder	3/23/2005 4:45 PM	
drivers		File Folder	3/23/2005 10:24 AM	R
idwlog		File Folder	3/23/2005 4:46 PM	
privates		File Folder	3/17/2005 6:51 PM	R
Program Files		File Folder	3/23/2005 4:31 PM	R
Program Files (x86)		File Folder	3/23/2005 4:32 PM	R
RECYCLER		File Folder	3/24/2005 5:08 PM	HS
serveradministrator		File Folder	3/23/2005 4:46 PM	R
System Volume Information		File Folder	3/23/2005 4:45 PM	HS
WINDOWS		File Folder	3/24/2005 4:37 PM	
wmpub		File Folder	3/23/2005 4:34 PM	
AUTOEXEC.BAT	0 KB	Windows Batch File	3/23/2005 4:34 PM	HSA
boot.ini	1 KB	Configuration Settings	3/23/2005 4:27 PM	HS
CONFIG.SYS	0 KB	System file	3/23/2005 4:34 PM	HSA
IO.SYS	0 KB	System file	3/23/2005 4:34 PM	RHSA
MSDOS.SYS	0 KB	System file	3/23/2005 4:34 PM	RHSA
ntdetect.com	47 KB	Application	2/28/2005 6:00 AM	RHSA
ntldr	320 KB	System file	2/28/2005 6:00 AM	RHSA
pagefile.sys	2,095,104 KB	System file	3/24/2005 4:40 PM	HSA
SIPOBJ.DBG	6 KB	DBG File	3/23/2005 4:31 PM	A

Figure 7. Windows Server 2003 x64 Editions Program Files folder for 64-bit program files and Program Files (x86) folder for 32-bit program files

[13] An important exception is in the case of data buffers passed to device drivers (the DeviceIoControl function). For more information, visit msdn.microsoft.com/library/default.asp?url=/library/en-us/kmarch/hh/kmarch/Other_6cff9ff8-fb41-4cb9-bbfd-b68e5fd17496.xml.asp.

[14] For more information about the file system redirector, visit msdn.microsoft.com/library/default.asp?url=/library/en-us/win64/win64/file_system_redirector.asp. By default, the WOW64 file system redirector is turned on. Programmers can enable or disable this feature by using the Wow64EnableWow64FsRedirection() application programming interface (API) that is available on x64. For more information, visit msdn.microsoft.com/library/default.asp?url=/library/en-us/sysinfo/base/wow64enablewow64fsredirection.asp.



Name	Size	Type	Date Modified	Attributes
Provisioning		File Folder	3/23/2005 10:16 AM	
Registration		File Folder	3/23/2005 4:30 PM	
repair		File Folder	3/23/2005 10:16 AM	
Resources		File Folder	3/23/2005 10:16 AM	
security		File Folder	3/23/2005 4:41 PM	
SoftwareDistribution		File Folder	3/23/2005 4:45 PM	
srchasst		File Folder	3/23/2005 4:30 PM	
system		File Folder	3/23/2005 10:16 AM	
system32		File Folder	3/24/2005 4:45 PM	
SysWOW64		File Folder	3/23/2005 4:34 PM	
TAPI		File Folder	3/23/2005 10:23 AM	
Tasks		Folder	3/23/2005 4:43 PM	S
Temp		File Folder	3/23/2005 5:12 PM	
twain_32		File Folder	3/23/2005 10:16 AM	
Web		File Folder	3/23/2005 4:32 PM	R
WinSxS		File Folder	3/23/2005 4:41 PM	
0.log	0 KB	Text Document	3/24/2005 4:41 PM	A
Blue Lace 16.bmp	2 KB	Bitmap Image	2/28/2005 6:00 AM	A
bootstat.dat	2 KB	DAT File	3/24/2005 4:40 PM	SA
certmmc.log	1 KB	Text Document	3/23/2005 4:33 PM	A
certocm.log	5 KB	Text Document	3/23/2005 4:41 PM	A
cmsetacl.log	1 KB	Text Document	3/23/2005 4:28 PM	A
Coffee Bean.bmp	17 KB	Bitmap Image	2/28/2005 6:00 AM	A
comsetup.log	21 KB	Text Document	3/23/2005 4:41 PM	A
control.ini	0 KB	Configuration Settings	3/23/2005 4:34 PM	A
desktop.ini	1 KB	Configuration Settings	2/28/2005 6:00 AM	A
dialer.exe	71 KB	Application	2/28/2005 6:00 AM	A
DtcInstall.log	2 KB	Text Document	3/23/2005 4:29 PM	A
explorer	1 KB	Windows Explorer C...	2/28/2005 6:00 AM	A
explorer.exe	1,367 KB	Application	2/28/2005 6:00 AM	A
FaxSetup.log	15 KB	Text Document	3/23/2005 4:30 PM	A
FeatherTexture.bmp	17 KB	Bitmap Image	2/28/2005 6:00 AM	A
Gone Fishing.bmp	17 KB	Bitmap Image	2/28/2005 6:00 AM	A
Greenstone.bmp	26 KB	Bitmap Image	2/28/2005 6:00 AM	A
hh.exe	12 KB	Application	2/28/2005 6:00 AM	A
iis6.log	61 KB	Text Document	3/23/2005 4:41 PM	A
imsins.log	11 KB	Text Document	3/23/2005 4:41 PM	A

Figure 8. Windows Server 2003 x64 Editions System32 folder for 64-bit system files and SysWOW64 folder for 32-bit system files

subdirectories in \Windows\System32 are excluded from automatic redirection by the file system redirector:

- %windir%\system32\drivers\etc
- %windir%\system32\spool
- %windir%\system32\catroot2
- %windir%\system32\logfiles

To accommodate 32-bit and 64-bit applications, the x64 registry layout also has been designed to ensure that hard-coded DLL paths, application settings, and other parameter values are not overwritten. The 32-bit and 64-bit processes in an x64 OS use different registry sections: native 64-bit processes use HKLM\Software, while 32-bit processes running in WOW64 mode access HKLM\Software\WOW6432Node. This process of automatically directing applications to the appropriate registry keys, which depends on whether the application is 32-bit or 64-bit, is implemented by WOW64 and is known as registry redirection.[15]

While 32-bit applications running on an x64 OS may experience nominal performance benefits because of the architectural differences in the 64-bit OS and x64-capable hardware, 32-bit applications should be upgraded to 64-bit versions to take full advantage of x64-capable hardware and OS capabilities. Figure 9 provides a list of Microsoft applications and their support status on Windows Server 2003 x64 Editions at press time. Please visit www.microsoft.com for the latest support information.

Applications supported in Compatibility mode (via WOW64)
.NET Framework 1.1 Visual Studio® .NET 2003 SQL Server 2000 SP4
Applications that are already 64-bit
Internet Information Services on Windows Server 2003 x64 Editions Microsoft Cluster Service (MSCS) for Windows Server 2003 x64 Editions Windows System Resource Manager (WSRM) shipping with Windows Server 2003 x64 Editions
Applications not supported on Windows Server 2003 x64 Editions
Exchange Server 2003 and Exchange Server 2003 SP1 SQL Server 2000 SP3a or earlier Virtual Server 2005 with x64 as the host OS

Figure 9. Microsoft applications that currently support Windows Server 2003 x64 Editions

An initial step in migrating to 64-bit technology

Even though Windows Server 2003 x64 Editions do not have "SP1" in their product names, these operating systems are built from the Windows Server 2003 SP1 code-tree. For that reason, they include most updates, features, and security enhancements introduced in Windows Server 2003 SP1—including data execution prevention, the Security Configuration Wizard, Windows Firewall enhancements, and Server Balanced Processor Power and Performance. By supporting both 32-bit and 64-bit applications, Windows Server 2003 x64 Editions on Intel EM64T architecture can enable organizations to integrate 64-bit technology into their data centers and gradually migrate to this emerging platform without requiring a major overhaul of their IT infrastructure.

Acknowledgments

The authors would like to thank Darrell Gorter, Joe Swart, Mona Abouzied, Samer Arafeh, and Sandy Arthur at Microsoft for their invaluable input on this article.

Ranjith Purush is a systems engineer in the Server Operating Systems Engineering Group at Dell. He is currently leading the engineering effort for Windows Server 2003 x64 Editions. He has published several articles on a wide range of subjects, including virtualization software, system recovery, and server consolidation. Ranjith has an M.S. in Electrical and Computer Engineering from The University of Texas at Austin.

Chip Webb is a strategist in the Server Operating Systems Engineering Group at Dell. He leads the effort to integrate current and future versions of Microsoft Windows and Novell® NetWare® on Dell PowerEdge servers. He has extensive experience with OS development and system software. Chip has a B.S. from Michigan State University.

[15] For more information about the registry redirector and other features such as registry reflection, visit msdn.microsoft.com/library/default.asp?url=/library/en-us/win64/win64/registry_redirector.asp. By default, WOW64 registry reflection is turned on. Programmers may override this feature by using certain registry APIs. For more information, visit msdn.microsoft.com/library/default.asp?url=/library/en-us/win64/win64/accessing_an_alternate_registry_view.asp.

Deploying Microsoft Windows Server 2003 x64 Editions on Dell PowerEdge Servers

Microsoft® Windows Server™ 2003 x64 Editions are designed for servers that are based on 64-bit extended architecture. This article provides guidelines for installing Windows Server 2003 x64 Editions on Dell™ PowerEdge™ servers equipped with Intel® Extended Memory 64 Technology–based processors. The article also describes hardware and software components supported by the x64 operating systems.

BY RANJITH PURUSH AND SANDHYA SENAPATHI

Related Categories:

Dell OpenManage

Dell PowerEdge servers

Intel EM64T processors

Microsoft Windows

Microsoft Windows Server 2003 x64 Editions

Operating system (OS)

System deployment

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

Microsoft Windows Server 2003 x64 Editions—also referred to as x64—support 64-bit extended architecture–based servers. These operating systems are supported on eighth-generation and later Dell PowerEdge servers[1] that are equipped with Intel Extended Memory 64 Technology (EM64T). Figure 1 lists the Dell PowerEdge servers that currently support the x64 platform.

To help ensure compatibility on Windows Server 2003 x64 Editions, Dell worked extensively with Microsoft and other hardware and software partners, performing comprehensive tests across Dell's hardware and software products throughout the development of the x64 platform. In addition to performing tests in Dell labs, Dell engineers worked with several enterprises that participated in Microsoft's Technology Adoption Program. These enterprises received prerelease versions of device drivers, BIOS and firmware releases, and Dell OpenManage™ 4 software components. The Technology Adoption Program helped identify and address enterprise issues from the perspective of a production environment, which in turn enhanced Dell hardware and software product readiness for Windows Server 2003 x64 Editions.

This article discusses Windows Server 2003 x64 Editions and the Intel EM64T architecture, focusing in particular on the installation of x64 on EM64T-based Dell PowerEdge servers.[2] Besides explaining the recommended installation process for Windows Server 2003 x64 Editions, this article describes Dell OpenManage components, network components, and storage components supported by x64.

Dell support for Windows Server 2003 x64 Editions

Microsoft offers three editions of x64: Standard, Enterprise, and Datacenter.[3] Dell supports Standard x64 Edition and Enterprise x64 Edition, which are compared in Figure 2. Figure 3 lists Microsoft's required minimum and recommended minimum hardware configurations for supporting these operating systems.

Windows Server 2003 SP1 and x64

Windows Server 2003 x64 Editions are based on the same code-tree as Windows Server 2003 Service Pack 1 (SP1). However, x64 is not packaged as a Service Pack. The x64 Editions are 64-bit operating systems, and they differ

[1] Legacy Dell PowerEdge servers (seventh-generation and earlier, not listed in Figure 1) do not support Intel EM64T–capable processors and thus do not support Windows Server 2003 x64 Editions. For more information about eighth-generation Dell PowerEdge servers that do support Intel EM64T architecture, visit www.dell.com/servers.

[2] For more information about Windows Server 2003 x64 Editions, see "Introducing Microsoft Windows Server 2003 x64 Editions for the Intel EM64T Platform" by Ranjith Purush and Chip Webb in *Dell Power Solutions,* May 2005.

[3] For a detailed product overview of Windows Server 2003 x64 Editions, visit www.microsoft.com/windowsserver2003/64bit/x64/default.mspx.

Dell PowerEdge server	BIOS/BMC firmware
PowerEdge SC1420	A00/None
PowerEdge SC1425	A01/A01
PowerEdge 1800	A02/A01
PowerEdge 1850	A02/A02
PowerEdge 1855	A02/A00
PowerEdge 2800	A02/A02
PowerEdge 2850	A02/A02
PowerEdge 6800	A00/A00
PowerEdge 6850	A00/A00

Note: To obtain the latest BIOS and BMC firmware, visit support.dell.com.

Figure 1. Minimum BIOS/BMC requirements for x64-capable Dell PowerEdge servers

from 32-bit versions of Windows Server 2003 as well as Windows Server 2003 for 64-Bit Itanium®-based Systems.

However, features and security enhancements introduced in 32-bit Windows Server 2003 SP1[4] are available on the x64 platform, and Dell PowerEdge servers based on Intel EM64T support features such as data execution prevention (DEP) and Server Balanced Processor Power and Performance. DEP requires processors that support Execute Disable (XD) as well as the minimum Dell system BIOS revision (see Figure 4). Dell PowerEdge servers shipped since October 2004 have XD-supported processors.[5]

Server Balanced Processor Power and Performance support in Windows Server 2003 x64 Editions is designed to leverage the Enhanced Intel SpeedStep® Technology[6] (EIST) on supported Intel processors. Support for this OS feature is dependent on the processor model, frequency,[7] and stepping. Figure 5 lists the minimum BIOS requirements to leverage EIST capabilities in supported Dell servers. To determine whether the processors[8] on a Dell PowerEdge server are EIST capable, administrators can take the following steps:

Edition	Physical memory	Physical processors	General features
Windows Server 2003, Standard x64 Edition	Up to 32 GB	Up to 4	Domain controller, Microsoft Active Directory® directory service, Internet Connection Firewall (ICF), Internet Authentication Service (IAS), Internet Connection Sharing (ICS), IPv6, Distributed File System (DFS), Windows Management Instrumentation (WMI), Internet Information Services (IIS) 6.0, Dynamic Host Configuration Protocol (DHCP), Domain Name System (DNS), Windows Internet Naming Service (WINS), Remote Installation Service, Terminal Server, and virtual private network
Windows Server 2003, Enterprise x64 Edition	Up to 1 TB	Up to 8	All the preceding features of Windows Server 2003, Standard x64 Edition, plus eight-node Microsoft Cluster Service (MSCS) support, Terminal Server with Session Directory, Metadirectory Services support, Remote Storage, non-uniform memory access (NUMA), and Windows System Resource Manager (WSRM)

Figure 2. Feature comparison between Standard x64 Edition and Enterprise x64 Edition

Supported hardware	Required minimum	Recommended minimum
Intel Xeon™ with Intel EM64T	2.80 GHz	3.60 GHz
Intel Pentium® with Intel EM64T	3.20 GHz	3.60 GHz
Memory	512 MB	512 MB
Disk space	4 GB*	4 GB*

** Dell recommends a 12 GB system partition. The default setting for Dell factory-installed images and the Dell PowerEdge Installation and Server Management CD is 12 GB.*

Figure 3. Hardware requirements for Windows Server 2003 x64 Editions

1. When the system is booting up, press F2 to enter system BIOS settings.
2. Under "CPU Information," check for the Demand-Based Power Management option.
3. If the Demand-Based Power Management option is available and editable as "Disabled" or "Enabled," all the processors on the server support EIST.
4. If the Demand-Based Power Management option is available and read-only, then at least one of the processors on the system does not support EIST.

Dell-supported peripherals for x64

Windows Server 2003 x64 Editions require that all drivers, including device drivers, be 64-bit drivers. Device drivers also need x64-specific decorators in the driver information (.inf) files.[9] Dell worked with peripheral device vendors as well as Microsoft to help ensure availability of 64-bit drivers for peripheral devices supported on EM64T-based Dell PowerEdge servers. As a result, many of these drivers have native[10] support in x64.

The following sections provide information about driver support for major peripheral devices such as network and storage adapters. The information includes:

- Minimum supported driver and firmware versions
- Whether specific peripherals have a native driver in the OS[11]
- PowerEdge servers on which specific peripherals are supported

[4] For more information about the changes introduced by SP1, see "Guide to Deploying Microsoft Windows Server 2003 Service Pack 1 on Dell PowerEdge Servers" by Min-John Lee, Scott M. Callaway, and Jeff Ferris in *Dell Power Solutions,* May 2005.

[5] For more information about XD, visit www.intel.com/business/bss/infrastructure/security/xdbit.htm.

[6] For more information about Enhanced Intel SpeedStep technology, visit www.intel.com/cd/ids/developer/asmo-na/eng/195910.htm.

[7] The minimum frequency requirement for EIST support varies with the processor model. For example, the current minimum frequency requirement for the Intel Nocona Xeon processor is 3.4 GHz.

[8] The Windows OS will not enable Server Balanced Processor Power and Performance unless all processors on the server are EIST capable.

[9] For more information about .inf file requirements that affect device installation on x64, visit www.microsoft.com/whdc/driver/install/64INF_reqs.mspx.

[10] Even if a device has a native driver in the OS, Dell recommends that administrators check for the latest drivers on support.dell.com.

[11] Drivers that are included in the OS are listed as "native" in the location column; drivers that are not included in the OS are listed as "non-native." Administrators can download non-native drivers and utilities from the Dell Web site (support.dell.com) or use the Dell PowerEdge Service and Diagnostic Utilities CD that shipped with the Dell server.

- Categorization of devices based on supported technology—that is, Peripheral Component Interconnect Extended (PCI-X) or PCI Express[12]

Dell PowerEdge server	Minimum BIOS revision required for XD support
PowerEdge SC1420	A00
PowerEdge SC1425	A00
PowerEdge 1800	A01
PowerEdge 1850	A02
PowerEdge 1855	A02
PowerEdge 2800	A02
PowerEdge 2850	A02
PowerEdge 6800	A00
PowerEdge 6850	A00

Figure 4. Minimum BIOS requirements for XD support on Dell PowerEdge servers

Dell PowerEdge server	Minimum BIOS revision required for EIST support
PowerEdge SC1420	EIST not supported
PowerEdge SC1425	A01
PowerEdge 1800	A01
PowerEdge 1850	A02
PowerEdge 1855	A02
PowerEdge 2800	A02
PowerEdge 2850	A02
PowerEdge 6800	A00
PowerEdge 6850	A00

Figure 5. Minimum BIOS requirements for EIST support on Dell PowerEdge servers

Note: The driver versions listed in this article were the minimum versions required to support Windows Server 2003 x64 Editions at press time. Dell recommends that administrators use the latest drivers, which can be found at support.dell.com.

Supported network components

Dell worked closely with two major vendors, Broadcom and Intel, to help ensure support for both embedded and add-on network peripherals on PowerEdge servers that support x64. Figure 6 provides additional information about x64 drivers for Intel and Broadcom network adapters supported on x64-capable Dell PowerEdge servers.

Advanced networking features: Teaming and bridging

Teaming enables a group of adapters to be configured together for various purposes such as increased throughput, load balancing, and fault tolerance. Teaming requires specialized software such as the Intel PROSet utility and the Broadcom Advanced Control Suite (BACS). The advanced intermediate drivers and base drivers required for teaming support are also part of these specialized software packages.

Broadcom and Intel have made available versions of the BACS and PROSet utilities that are compatible with x64. Figure 7 lists the minimum driver versions of these utilities that support x64. BACS is a 32-bit application suite that can execute on both 32-bit and x64 versions of Windows Server 2003, but with different kernel-mode drivers. The Intel PROSet utility has two separate versions for 32-bit and x64 primarily because of changes in the names of the installation files. The latest versions of both BACS and PROSet are available on the Dell Web site at support.dell.com.

Bridging is also supported natively in Windows Server 2003 x64 Editions. A network bridge is designed to create connections between different types of network media, helping administrators to manage LAN segments and to create a single subnet for the entire network. Additional hardware devices, drivers, or software are not required for bridging on the x64 platform.

Product	Driver/version/location	Supported PowerEdge servers	Connection speed	Technology
Add-on network adapters				
Intel PRO/100 S	efe5b32e.sys/7.1.8.4/native	PowerEdge SC1420 PowerEdge SC1425 PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850	Fast Ethernet	PCI
Intel PRO/1000 MT	e1G5132e.sys/8.1.4.0/native		Gigabit Ethernet	PCI-X
Intel PRO/1000 MT Dual Port	e1G5132e.sys/8.1.4.0/native		Gigabit Ethernet	PCI-X
Intel PRO/1000 MF	e1G5132e.sys/8.1.4.0/native	PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850	Gigabit Ethernet	PCI-X
Intel PRO/1000 P Dual Port	e1G5132e.sys/8.4.21.0/non-native	PowerEdge SC1420 PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850	Gigabit Ethernet	PCI Express
Broadcom 5721	b57amd64.sys/7.107/non-native	PowerEdge SC1420 PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850	Gigabit Ethernet	PCI Express
Embedded network adapters				
Intel Gigabit Ethernet adapters	e1G5132e.sys/8.1.4.0/native	PowerEdge SC1420 PowerEdge SC1425 PowerEdge 1800 PowerEdge 1850 PowerEdge 1855 PowerEdge 2800 PowerEdge 2850	Gigabit Ethernet	PCI-X
Broadcom 5704	b57amd64.sys/7.80.0.0/native	PowerEdge 6800 PowerEdge 6850	Gigabit Ethernet	PCI-X

Figure 6. Minimum supported x64 driver revisions for embedded and add-on network peripherals

Adapter teaming utility	Driver/version/location
Intel PROSet	iansw32e.sys/8.01.05/non-native
Broadcom Advanced Control Suite	basp.sys/6.1.6/non-native

Figure 7. Teaming support on Windows Server 2003 x64 Editions

[12] For more information about Windows support for PCI Express technology, visit www.microsoft.com/whdc/system/bus/PCI/PCIe_Windows.mspx.

Supported storage components

Dell worked closely with two major storage vendors, LSI Logic and Adaptec, to help ensure support for both embedded and add-on storage adapters on PowerEdge servers that support x64. LSI Logic and Adaptec are both original equipment manufacturers (OEMs) for Dell PowerEdge Expandable RAID Controllers (PERCs) and Cost Effective RAID Controllers (CERCs). Figure 8 provides information about x64 drivers for supported storage controllers. Windows Server 2003 x64 Editions also offer OS-based software RAID.[13]

Eighth-generation Dell PowerEdge servers offer both SCSI and Serial ATA (SATA) RAID adapters. Critical storage devices that do not have native drivers in Windows Server 2003 x64 Editions include SCSI-based storage devices—the Adaptec 39320 and the PERC 320, Dual Channel (PERC 320/DC)—and SATA-based storage devices—CERC SATA 2s and CERC SATA 6ch.

Native backup support and supported secondary storage components

Windows Server 2003 x64 Editions support NTBackup,[14] a backup utility native in Windows Server 2003 that can be configured to help protect data from accidental loss if a system experiences hardware or storage media failure. NTBackup in x64 is a native 64-bit application and can be used to create a duplicate copy of data on a backup storage medium such as a hard drive, a removable disk, or an entire library of disks or tapes.

Secondary storage systems from Dell include Dell PowerVault™ stand-alone tape drives, tape autoloaders, and tape libraries. Figures 9 and 10 list the x64 drivers for these supported storage components.

Windows Server 2003 x64 Editions also support the native CD-burning software tool that allows data to be written to CD-R and CD-RW discs. This tool is implemented by the native Image Mastering Application Programming Interface (IMAPI) CD-burning COM Service.[15] It has a limited feature set and is not supported by NTBackup to back up or restore files.

Product	Driver/version/location	Supported PowerEdge servers	Technology	Firmware version
Add-on SCSI RAID adapters				
PERC 4/SC	mraid35x.sys/6.37.2.64/native	PowerEdge 1800 PowerEdge 1850	PCI-X	351H
PERC 4/DC	mraid35x.sys/6.37.2.64/native	PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850	PCI-X	351H
PERC 4e/DC	mraid35x.sys/6.37.2.64/native	PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850	PCI Express	521H
PERC 320/DC	aac.sys/4.0.0.5815/non-native	PowerEdge SC1420 PowerEdge 1800	PCI-X	5813
Adaptec 39320 with host software RAID*	A320raid.sys/2.00.00.76/ non-native	PowerEdge SC1420 PowerEdge SC1425 PowerEdge 1800	PCI-X	Not applicable
Embedded SCSI RAID adapters				
PERC 4e/Si	mraid35x.sys/6.44.3.64/native	PowerEdge 1850	PCI Express	521H
PERC 4e/Di	mraid35x.sys/6.44.3.64/native	PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850	PCI Express	521H
Add-on SATA RAID adapters				
CERC SATA 6ch	cercsr6.sys/4.1.1.7033/ non-native	PowerEdge SC1420 PowerEdge SC1425 PowerEdge 1800	PCI-X	4.1.0.7403
Embedded SATA RAID adapters				
CERC SATA 2s with host software RAID*	aarich.sys/6.00.00.076/ non-native	PowerEdge SC1420 PowerEdge SC1425 PowerEdge 1800	PCI-X	Not applicable
Add-on SCSI adapters				
Adaptec 39160	adpu160m.sys/RTC_XP107/ native	PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850	PCI-X	2.57.2s1
Embedded SCSI adapters				
LSI Logic 1020/1030	symmpi.sys/1.09.11.52/native	PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850	PCI-X	5.06.04

** The Adaptec 39320 SCSI RAID adapter—also known as the Adaptec U320 SCSI RAID 0 or 1 controller—and the CERC SATA 2s adapter provide driver-based RAID with RAID-0 and RAID-1 capabilities.*

Figure 8. Minimum supported x64 driver revisions for embedded and add-on storage controllers

Installation process for Windows Server 2003 x64 Editions

Because Windows Server 2003 x64 Editions differ significantly from other 32-bit and 64-bit Microsoft offerings, they require a

[13] For more information about OS-based software RAID, visit www.microsoft.com/technet/prodtechnol/windowsserver2003/library/ServerHelp/1279a8c6-1b47-482d-bc00-4b4f91ec6412.mspx.

[14] For more information about native backup support available on Windows Server 2003, visit www.microsoft.com/technet/prodtechnol/windowsserver2003/library/ServerHelp/7803d7f2-390c-42fe-9171-c825c4b11668.mspx.

[15] For more information about the native Windows Server 2003 x64 CD-burning software utility, visit support.microsoft.com/default.aspx?scid=kb;en-us;317525.

Product	Driver/version/location	Supported PowerEdge servers
PowerVault 100T Travan40	qic157.sys/5.2.3790/native	PowerEdge SC1420 PowerEdge 1800
PowerVault 100T DDS4	4mmdat.sys/5.2.3790/native	PowerEdge 6800 PowerEdge 6850
PowerVault 100T DAT72	Pvdatw2k.sys/1.11.0.0/ non-native	PowerEdge SC1420 PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850
PowerVault 110T DLT VS 80	dlttape.sys/5.2.3790/native	PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850
PowerVault 110T LTO-1	ltotape.sys/5.2.3790/native	PowerEdge 1800 PowerEdge 1850 PowerEdge 2800 PowerEdge 2850 PowerEdge 6800 PowerEdge 6850
PowerVault 110T LTO-2	dtapex64.sys/6.0.68/non-native	
PowerVault 110T LTO-2-L	pvlto.sys/1.9.0.0/non-native	
PowerVault 110T LTO-3	dtapex64.sys/6.0.68/non-native	
PowerVault 110T DLT VS 160	Qdltx64.sys/3.0.1.0/non-native	
PowerVault 110T SDLT 320	qdltx64.sys/3.0.1.0/non-native	

Figure 9. Minimum supported x64 driver revisions for stand-alone PowerVault tape drives

fresh installation (see Figure 11).[16] However, the installation process for Windows Server 2003 x64 Editions on supported Dell PowerEdge servers is similar to that of the 32-bit version of Windows Server 2003. Dell supports three ways to install Windows Server 2003 x64 Editions: clean installation using the Windows Server 2003 x64 CD, clean installation using the Dell PowerEdge Installation and Server Management CD, and Dell factory installation.[17]

Windows Server 2003 x64 CD

For a clean installation of Windows Server 2003 x64 Editions on Dell PowerEdge servers, administrators should perform the following steps sequentially:

1. **Prepare BIOS and firmware.** A clean installation begins with an upgrade of the Dell server's BIOS, baseboard management controller (BMC) firmware, and primary storage controller firmware. Before installing x64, administrators should ensure that they have the appropriate BIOS, BMC firmware, and primary storage controller firmware as discussed in this article. Administrators should either download these components from the Dell Web site (support.dell.com) or use the latest Dell PowerEdge Service and Diagnostic Utilities CD that supports Windows Server 2003 x64 Editions (see the "Dell OpenManage support for Windows Server 2003 x64 Editions" section in this article for more information). Depending on the server and storage adapters on the server, an F6 installation[18] may be required to ensure that the device drivers for the boot drives are loaded. See the "Supported storage components" section in this article for more information about which devices can be loaded with the native drivers and which devices require non-native third-party drivers.
2. **Install the x64 OS.** Administrators can perform a clean installation using the Windows Server 2003 x64 OS installation CD.

Product	Supported drives	Driver/version/location	Supported PowerEdge servers
PowerVault 112T enclosure	DDS4/DLT VS 80/DLT VS 160	Depends on the tape drive carried	PowerEdge 1850, PowerEdge 2800, PowerEdge 2850, PowerEdge 6800, PowerEdge 6850
PowerVault 114T enclosure	DAT72/DLT VS 160/SLDT 320/LTO-1/LTO-2/ LTO-2-L/LTO-3	Depends on the tape drive carried	PowerEdge 1800, PowerEdge 1850, PowerEdge 2800, PowerEdge 2850, PowerEdge 6800, PowerEdge 6850
PowerVault 122T autoloader	DLT VS 80/LTO-1/LTO-2/SDLT 320	powerfil.sys/5.2.3790/native* (if the tape drive carried has native support)	PowerEdge 1800, PowerEdge 1850, PowerEdge 2800, PowerEdge 2850, PowerEdge 6800, PowerEdge 6850
PowerVault 132T library	LTO-2/LTO-3/SDLT 320	pv132t.sys/6.0.0.0/non-native*	
PowerVault 136T library	LTO-1/LTO-2/SDLT 320	adicsc.sys/5.2.3790/native* (if the tape drive carried has native support) pv136t.sys/6.0.0.0/non-native* (if the tape drive carried does not have native support)	PowerEdge 1850, PowerEdge 2800, PowerEdge 2850, PowerEdge 6800, PowerEdge 6850

** This information pertains to the driver/version/location of the autoloader or library; the driver/version/location of the tape drive depends on the particular tape drive being used.*

Figure 10. Minimum supported x64 driver revisions for PowerVault enclosures, autoloaders, and libraries

[16] No upgrade path is available to Windows Server 2003 x64 Editions from any 32-bit or 64-bit Windows OS. Attempting to run the x64 version of the WINNT32 setup program in a 32-bit OS environment with the intention of upgrading from a 32-bit platform to x64 will result in the "WINNT32.exe is not a valid Win32 application" informational message. However, an upgrade path is available from Windows Server 2003, Standard x64 Edition, to Windows Server 2003, Enterprise x64 Edition.

[17] Alternate solutions from Microsoft for installing Windows Server 2003 x64 Editions include Unattended installation using unattend.txt and winnt.sif; System Preparation Tool (Sysprep) image deployment; and Remote Installation Service (RIS) deployment. *Note:* To deploy x64 using RIS, the RIS server must be running Windows Server 2003 SP1 or later.

[18] *F6 installation* refers to pressing F6 during the OS installation to install third-party storage drivers for devices that do not have native drivers on the Windows Server 2003 x64 CD. Windows Server 2003 x64 Editions support adding non-native drivers via F6.

3. **Verify device drivers.** After installing the x64 OS, administrators must use the Device Manager to verify that device drivers have installed with no problems and are working correctly. A yellow exclamation point next to a device in Device Manager usually indicates that a driver is needed. Administrators can download the necessary drivers from the Dell Web site (support.dell.com) or from the Dell PowerEdge Service and Diagnostic Utilities CD.
4. **Install x64 versions of Dell tools and software.** Administrators complete the installation process by installing x64 versions of Dell tools and software, such as the Dell OpenManage components. See the "Dell OpenManage support for Windows Server 2003 x64 Editions" section in this article for more information.

Microsoft supports dual-boot configuration that includes Windows Server 2003 x64 Editions and 32-bit Windows, allowing administrators to boot into either OS. However, to help protect against potential application incompatibility issues, data loss, or system instability in a dual-boot configuration, Microsoft recommends that administrators install the 32-bit OS and x64 on separate partitions and that they install the 32-bit OS before installing x64.

Dell PowerEdge Installation and Server Management CD

The Dell PowerEdge Installation and Server Management CD[19] guides administrators—via easy-to-use graphical user interface (GUI) menus—through a clean OS installation on a Dell PowerEdge server. However, this method still requires using the Windows Server 2003 x64 CD for the OS software.

Dell strongly recommends that administrators use the Dell PowerEdge Installation and Server Management CD for the OS installation process when the factory installation option is not chosen. In addition to simplifying the installation process with easy-to-use GUI menus, the Dell CD provides the latest drivers for all supported devices on Dell PowerEdge servers.

The Dell OpenManage CD kit that ships with Dell servers includes the Dell PowerEdge Installation and Server Management CD. This kit also contains the Dell PowerEdge Service and Diagnostic Utilities CD that contains drivers and utilities required to support Windows Server 2003 x64 Editions on PowerEdge servers.

Dell factory installation

Dell plans to offer factory installation for Windows Server 2003 x64 Editions later in 2005. This option will allow enterprises to have the OS installed during the Dell server manufacturing process. A server purchased from Dell with x64 pre-installed at the factory will have the latest available device drivers and firmware.

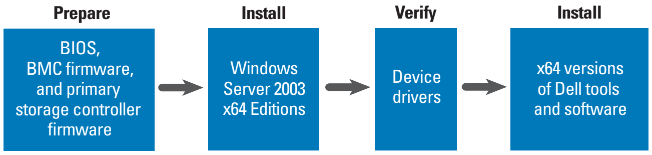

Prepare
BIOS, BMC firmware, and primary storage controller firmware
Install
Windows Server 2003 x64 Editions
Verify
Device drivers
Install
x64 versions of Dell tools and software

Figure 11. The clean installation process for Windows Server 2003 x64 Editions

Dell OpenManage support for Windows Server 2003 x64 Editions

The Dell OpenManage infrastructure is a systems management application suite that offers proactive monitoring, diagnostics, notification, and remote access for Dell PowerEdge servers. Dell OpenManage 4.4 and later will support Windows Server 2003 x64 Editions. The Dell OpenManage application suite for Windows Server 2003 x64 Editions will continue to be 32-bit (running in WOW64 mode) except for the 64-bit kernel-mode drivers. Dell OpenManage 4.4 will include Dell OpenManage Storage Services (OMSS), which is the management utility for storage components. OMSS will be the only[20] storage management software from Dell that supports Windows Server 2003 x64 Editions.

A 64-bit platform that supports 32-bit systems

By following the guidelines and procedures outlined in this article, administrators can deploy Microsoft Windows Server 2003 x64 Editions on Intel EM64T–based Dell PowerEdge servers. The 64-bit extended architecture of these servers and the support for both 32-bit and 64-bit applications provided by Windows Server 2003 x64 Editions can enable organizations to begin the migration to 64-bit platforms while still using their existing 32-bit systems. The coexistence of 32-bit and 64-bit technology enables organizations to maximize their IT investments while preparing for the future.

Acknowledgments

The authors would like to thank their colleagues in the Server Operating Systems Engineering Group as well as other engineering teams at Dell for their invaluable input on this article.

Ranjith Purush is a systems engineer in the Server Operating Systems Engineering Group at Dell. He is currently leading the engineering effort for Windows Server 2003 x64 Editions. Ranjith has an M.S. in Electrical and Computer Engineering from The University of Texas at Austin.

Sandhya Senapathi is a systems engineer in the Server Operating Systems Engineering Group at Dell. Sandhya has an M.S. in Computer Science from The Ohio State University.

[19] Refer to the Dell PowerEdge Documentation CD for more information about the contents of the Dell PowerEdge Installation and Server Management CD and installation best practices.

[20] The legacy Dell OpenManage Array Manager storage management tool is nearing end of life and thus will not be supported on Windows Server 2003 x64 Editions.

Guide to Deploying Microsoft Windows Server 2003 Service Pack 1 on Dell PowerEdge Servers

Microsoft® Windows Server™ 2003 Service Pack 1 (SP1) incorporates a set of security enhancements and tools designed to help administrators more effectively manage the security of their server installations when upgrading to SP1 on Windows Server 2003 systems or installing Windows Server 2003 with SP1 integrated. This article provides recommendations on the deployment process for Dell™ PowerEdge™ servers and discusses the key security features and remote management changes implemented in Windows Server 2003 SP1.

BY MIN-JOHN LEE, SCOTT M. CALLAWAY, AND JEFF FERRIS

Related Categories:

Change management

Dell OpenManage

Dell PowerEdge servers

Microsoft Windows Server 2003

Microsoft Windows Operating system (OS)

Remote management

Security

System deployment

Systems management

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

Deploying Microsoft Windows Server 2003 Service Pack 1 (SP1) can help enhance security and reliability, and simplify administrative tasks in environments using systems such as the Dell PowerEdge 1850, PowerEdge 2850, PowerEdge 6650, and PowerEdge 6850 servers as well as the PowerEdge 1855 blade server. Windows Server 2003 SP1 is the first cumulative service pack upgrade for the Windows Server System™ 2003 release. Although many of the security enhancements in SP1 have already been introduced in Microsoft Windows® XP Service Pack 2 (SP2) for the client environment, the server environment is characterized by specific traits that necessitated the SP1 release for Windows Server 2003. SP1 introduces certain features that require hardware-level support in the server, including data execution prevention (DEP) and demand-based switching (DBS).

Dell and Microsoft engineers worked together closely to support holistic SP1 software and hardware development, and performed extensive testing across supported Dell PowerEdge servers and Dell PowerVault™ network attached storage (NAS) servers to help ensure the compatibility and stability of Dell software and hardware. In addition, Dell plans to release version 4.4 of the Dell OpenManage™ infrastructure in May 2005 to support the security enhancements and features in Windows Server 2003 SP1.

Dell supports Windows Server 2003 SP1 on server platforms that support the original Windows Server 2003 release—including third-generation through seventh-generation Dell PowerEdge servers as well as eighth-generation PowerEdge servers. This article is intended to help guide administrators in deploying SP1 on Dell PowerEdge servers and PowerVault NAS servers by examining two deployment scenarios: upgrading to SP1 on existing Windows Server 2003 systems and installing Windows Server 2003 with SP1 integrated.

In addition, this article addresses application compatibility and server manageability issues relating to the following major technologies in SP1:

- The DEP feature
- Windows Firewall
- Remote systems management

Best practices for SP1 deployment

The first step in any deployment process is a careful evaluation of the existing IT environment. Documenting infrastructure—such as system BIOS, system and device firmware, and device driver versions; applications; and network components—is key to a successful service pack upgrade. In addition, administrators must first back up critical data and check systems for spyware and other unwanted software before upgrading to another service pack.

Performing essential housecleaning before deployment also helps smooth the migration process. Administrators should always perform BIOS, firmware, and driver updates prior to an OS upgrade.[1] The latest BIOS, firmware, and drivers are available from the Dell Web site or the Dell OpenManage management suite.

Besides updating BIOS, firmware, and drivers, administrators should check application compatibility before deploying any service pack. For an application compatibility evaluation, administrators can visit the Microsoft Windows Application Compatibility Web site and download the latest Application Compatibility Toolkit.[2]

Deployment path for upgrading to SP1 on existing Windows Server 2003 systems

Before proceeding with deployment, administrators should note that specific Dell PowerEdge hardware configurations with factory-installed Windows Server 2003 operating systems may have a registry issue with the Windows Server 2003 SP1 upgrade. Administrators should run the Dell Registry Preparation tool (regprep) for these configurations prior to upgrading to SP1. For more information about the regprep utility and which servers may require preparation, visit support.dell.com/support/topics/global.aspx/support/kb/en/document?c = us&cs = 555&DN = 1092292&l = en&s = biz. When upgrading current Windows Server 2003 systems to SP1, administrators have the following options:

- Upgrade from local media using the SP1 installation CD
- Install from a network share containing the installation files
- Upgrade over the Internet using Microsoft Windows Update[3]
- Automate the deployment process by using an enterprise software deployment tool such as Microsoft Systems Management Server 2003 (SMS 2003)

Upgrading from local media is the simplest method of installing Windows Server 2003 SP1. Upgrading from a network share is also a simple installation method and eliminates the need for media.

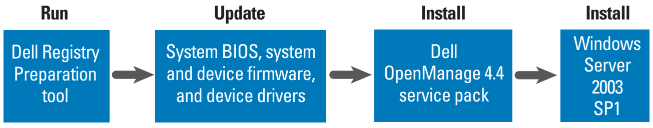

Run
Dell Registry Preparation tool
Update
System BIOS, system and device firmware, and device drivers
Install
Dell OpenManage 4.4 service pack
Install
Windows Server 2003 SP1

Figure 1. Recommended installation process on servers running Dell OpenManage 4.3

To use Microsoft Windows Update for SP1 deployment, administrators should go to the Windows Update Web site, install the update plug-in for Internet Explorer, and then install SP1. Service packs are listed in the High Priority Updates section. Administrators can configure updates to download automatically and then install applicable service packs and hot fixes either automatically or manually.

Each of the three preceding options—upgrading from local media, installing from a network share, and upgrading over the Internet using Windows Update—may entail a lengthy process for organizations that have many servers to upgrade. Thus, the fourth option—automating the process using an enterprise software deployment tool—is the preferred method for most large and midsize organizations. Many enterprise management tools exist; however, Microsoft SMS 2003 is designed to streamline SP1 upgrades with its integrated Distribute Software Updates Wizard. After authorizing Microsoft Windows Server 2003 SP1 in the SMS 2003 administration console, administrators can configure SMS 2003 to identify any systems joining the managed network and then deploy SP1 without manual intervention. Administrators can also configure SP1 settings by establishing group policies or by using an additional package distributed by SMS.[4]

To upgrade to SP1 on an existing system running Windows Server 2003, Dell supports the two following deployment paths:

- **Dell OpenManage 4.3:** Administrators should run the regprep tool;[5] update the system BIOS, system and device firmware, and device drivers; install the Dell OpenManage service pack for version 4.4 (which will be available at support.dell.com); and then install SP1 (see Figure 1).

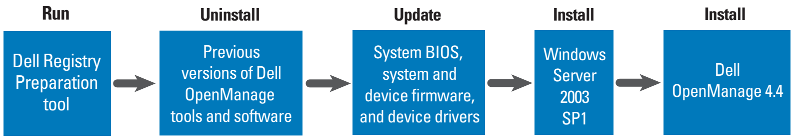

Run
Dell Registry Preparation tool
Uninstall
Previous versions of Dell OpenManage tools and software
Update
System BIOS, system and device firmware, and device drivers
Install
Windows Server 2003 SP1
Install
Dell OpenManage 4.4

Figure 2. Recommended installation process on servers running Dell OpenManage 4.2 or earlier

[1] For information about BIOS, firmware, and driver updates on specific Dell PowerEdge and PowerVault NAS servers, visit support.dell.com.

[2] For information about application compatibility, visit www.microsoft.com/windows/appcompatibility/default.mspx. To download the Microsoft Windows Application Compatibility Toolkit, visit msdn.microsoft.com/library/en-us/dnanchor/html/appcompat.asp.

[3] For Windows Update information and downloads, visit windowsupdate.microsoft.com.

[4] For more information about incorporating SMS 2003 into a deployment strategy, visit www.microsoft.com/smserver.

[5] For more information about regprep use, visit support.dell.com and search on the keyword "regprep."

- **Dell OpenManage 4.2 or earlier:** Administrators should run regprep; uninstall the previously installed Dell OpenManage tools and software; update the system BIOS, system and device firmware, and device drivers; install SP1; and then install Dell OpenManage 4.4 (see Figure 2).

For the latest SP1 upgrade information or compatibility alert, visit www.dell.com/microsoft.

Deployment path for installing Windows Server 2003 with SP1 integrated

For a new deployment of Windows Server 2003 with SP1 integrated, Dell offers several methods for ordering and installing Microsoft operating systems on Dell PowerEdge servers:

- Dell factory installation
- Dell OpenManage Server Assistant 8.6
- Dell Professional Services
- Altiris Deployment Solution for Dell Servers
- Microsoft Windows provisioning methods

Dell factory installation. Dell engineers worked closely with Microsoft engineers to validate and incorporate the latest Dell-qualified and Microsoft-qualified drivers into preinstallation OS images. If organizations order a Dell PowerEdge server with the option to have Windows Server 2003 with SP1 preinstalled, Dell deploys a custom OS image when the system is built in the Dell manufacturing facility. This option is designed to ensure that purchased systems integrate the latest Dell BIOS, firmware, and drivers as well as the latest version of Dell OpenManage infrastructure.

Dell OpenManage Server Assistant 8.6. Bundled with Dell OpenManage 4.4, Dell OpenManage Server Assistant (DSA) 8.6 supports a clean OS installation of Windows Server 2003 with SP1 on Dell PowerEdge servers. The System Update Utility in the Dell OpenManage 4.4 release also contains system BIOS updates, firmware, drivers, and utilities that administrators require to deploy and manage PowerEdge servers. Dell includes DSA with PowerEdge servers and also makes DSA available through the Dell OpenManage Subscription Service.[6]

Dell Professional Services. Dell offers many fee-based custom solutions that can be tailored to help reduce the impact of server upgrades and deployments on the supporting IT organization.[7]

Altiris Deployment Solution for Dell Servers. Dell and Altiris have collaborated to provide a simple-to-use deployment solution called the Altiris Deployment Solution for Dell Servers.[8] This approach provides administrators with easy-to-modify deployment scripts that can be used to manage system deployment for Dell PowerEdge servers.

Microsoft Windows provisioning methods. Microsoft has designed the following four tools to help automate SP1 deployment and customize installations:[9]

- Microsoft System Preparation (sysprep.exe)
- Unattended Setup (winnt32.exe)
- Remote Installation Services (RIS)
- Automated Deployment Services (ADS)

Sysprep.exe, which is included with the Microsoft Windows OS, helps administrators perform image-based installations of identical operating systems and software configurations on multiple systems quickly and efficiently. For unattended installation, Microsoft offers several tools that use answer files to automate the installation process. Answer files enable administrators to quickly install the Microsoft OS in Unattended Setup mode on multiple servers. Because answer files contain the required setup information—including system name, network adapter configuration, and Windows Firewall configuration—they enable administrators to easily perform unattended installations of Microsoft operating systems on multiple servers.

RIS and ADS are designed to permit network-initiated setup, enabling administrators to deploy both client and server operating systems on bare-metal servers that support Preboot Execution Environment (PXE).[10] Starting in SP1, network-based OS deployment is more secure because, during OS installation, the OS installation program applies a lock-down policy to the network interface to help prevent network-based attacks from occurring before security settings have been configured.

Application compatibility and server management

The security enhancements, features, and changes in SP1 may lead to application compatibility and server manageability concerns. This section addresses compatibility and manageability issues for three main aspects of SP1: the DEP feature, Windows Firewall, and remote systems management. In addition, this section discusses

[6] For more information about the Dell OpenManage Subscription Service, visit www1.us.dell.com/content/topics/global.aspx/services/en/om_subscr_svc?c=us&cs=04&l=en&s=bsd.

[7] For more information about Dell services, visit www.dell.com/services or contact a Dell sales representative.

[8] For more information about systems management products from Dell and Altiris, visit www.dell.com/altiris and see "Simplifying IT Operations with Altiris Deployment Solution for Dell Servers" by Todd Muirhead; Dave Jaffe, Ph.D.; and Landon Hale in *Dell Power Solutions,* May 2005.

[9] For more information about Windows OS provisioning methods, see "Guide to Deploying Microsoft Windows Server 2003 on Dell PowerEdge Servers" by the Dell Server Operating Systems Engineering Group in *Dell Power Solutions,* Special Issue, May 2003.

[10] For a list of the operating systems that can be deployed using RIS or ADS and for a comparison of RIS and ADS, visit support.microsoft.com/?kbid=842564.

two security tools introduced in SP1 to help provide post-installation server security management:

- **Security Configuration Wizard:** This tool is designed to allow system administrators to easily create and deploy security policies.
- **Post-Setup Security Updates:** This tool is designed to allow the newly installed OS to safely connect to the Internet and perform security updates.

See the "Windows Firewall" and "Remote systems management" sections in this article for more information about these two post-installation server security management tools.

Data execution prevention

DEP describes a set of technologies that help protect against malicious exploits by using a combination of hardware- and software-enforced memory protection methods. Hardware DEP implementations are available for 32-bit platforms running Physical Address Extension (PAE) or 64-bit extended architecture. Hardware-based DEP requires no-execute (NX)–capable processors. Dell PowerEdge servers shipped since October 2004 have NX-capable processors.[11]

In hardware DEP implementations, the processor keeps track of virtual memory pages, determining on a per-page basis whether a memory page should contain executable code. If a page reserved for nonexecutable code attempts to execute code, the hardware catches the exception and prevents the code from running.

Software-enforced DEP under Windows Server 2003 SP1 augments hardware DEP by providing an additional layer of security checks to prevent potential malicious exploitation of the exception-handling mechanisms in Windows Server 2003. Software DEP works alone or with compatible microprocessors to mark memory locations as NX. If a program tries to run any code—malicious or not—from a protected NX memory location, DEP closes the program and notifies the administrator.

To support hardware DEP, the system processor must support NX technology, the system BIOS must be NX-aware, and required PAE modules must be loaded during OS boot. Because the default setting in SP1 is to turn on hardware and software DEP for both OS kernel services and application levels, it is critical that administrators evaluate driver and application compatibility before deploying SP1. Many 64-bit device drivers were written for 64-bit versions of Windows and were required to be DEP- and PAE-compliant to function properly. Administrators should use the Dell Software Update Utility CD to update device drivers before upgrading to Windows Server 2003 SP1. *Note:* On 32-bit Windows versions running on systems supporting hardware DEP, device drivers may encounter technical issues caused by DEP or PAE mode being enabled. However, Dell has performed extensive testing and Microsoft Windows Hardware Quality Labs (WHQL) qualification on all supported device drivers.

For application compatibility, software developers must explicitly define executable memory segments in their application code.[12] If a business application encounters a compatibility issue after upgrading to SP1, developers can add the application to the DEP application exception list until the issue is resolved. To access the DEP administrative page in the system applet, administrators can right-click on My Computer, select the Properties menu item, click the Advanced tab, select Settings from the Performance section, and click the DEP tab.

BIOS requirements for NX and DBS support. Because hardware DEP requires memory protection–capable processors, Dell servers equipped with NX-capable Intel® processors require a BIOS update. A BIOS update is also required to support DBS. By throttling down processor frequency when the OS determines the processor utilization rate is low, DBS can help save power. DBS support in the OS leverages Enhanced Intel SpeedStep® Technology[13] and is dependent on the processor model, frequency, and stepping. To determine whether a given Dell PowerEdge server supports DBS, administrators can check the CPU Information menu in the BIOS settings. If the Demand-Based Power Management option is editable, then all processors in the system support DBS. If the option is not editable, at least one processor in the system does not support DBS. To turn on the DBS feature in the OS, select the Power Options icon in the Control Panel, and then select the "Server Balanced Processor Power and Performance" power scheme.

Mitigation. For server systems engineers, many system-level DEP configuration options can be controlled using the `/noexecute=DEP_option` switch specified in the boot.ini file, where `DEP_option` can be one of the following:

- `OptIn`: DEP is enabled for Windows programs and system services, and for other applications that have been explicitly identified.
- `OptOut`: DEP is enabled for applications and services. Specific applications can be excluded from DEP using the DEP application exception list or using the Microsoft Application Compatibility Toolkit as a reference.
- `AlwaysOn`: DEP applies to processes, with no exceptions.
- `AlwaysOff`: DEP does not apply to processes, and the processes will not run in PAE mode unless the `/PAE` switch is specifically included in the boot.ini entry.

[11] For more information about NX-capable processors, visit www.intel.com/business/bss/infrastructure/security/xdbit.htm.

[12] For the most up-to-date application compatibility information, visit msdn.microsoft.com.

[13] For more information about Enhanced Intel SpeedStep Technology, visit www.intel.com/cd/ids/developer/asmo-na/eng/195910.htm.

For scripted deployments, the preceding DEP options can be specified through the unattend.txt file.

Windows Firewall

Windows Server 2003 SP1 is designed to enable the same firewall features for servers that Windows XP SP2 provides for desktop computers. The default firewall setting is "Off" after a clean installation of Windows Server 2003 with SP1 integrated.

For an SP1 upgrade, firewall settings honor the pre-SP1 configuration. If administrators enable the firewall after an SP1 upgrade, they must identify which applications and network ports are required for the servers in the environment to provide services to network clients. Administrators can add these applications and network ports to the firewall exception list, identify which network clients can access specific services or applications, and control exceptions independently for each network interface card (NIC) in the system.

Once administrators have identified necessary exceptions, they can configure firewall options on individual systems by selecting the Windows Firewall applet from the Control Panel or by using the `netsh` command from the command line. For example, the `netsh firewall set portopening TCP 3389 ENABLE` command allows connections to TCP port 3389—the default port for Windows Terminal Server and Remote Desktop for Administration. The configuration set using either the applet or command line will be persistent unless it conflicts with options configured through a domain group policy. In a Microsoft Active Directory® directory service domain environment, group policy can be used to enable or disable Windows Firewall and configure exceptions for groups of servers.

The Security Configuration Wizard (SCW) is a server-specific tool introduced in SP1 that allows system administrators to easily create a set of security policies based on the server role, and apply the security policy set to one server or a group of servers. A SCW security policy includes Windows Firewall configuration, configuration of the system registry, and turnoff of unused system services to reduce attack surface.[14]

Remote systems management

Post-Setup Security Updates (PSSU) is a feature introduced in SP1 that enables Windows Firewall services and runs automatically in the console session directly following a clean installation of Windows Server 2003 with SP1 integrated. The purpose of this feature is to allow a system to safely connect to the Internet and perform security updates. The default network security policy is to block incoming traffic on every network port except network ports required to perform PSSU over the Internet.

Because the network connection to the target system is blocked during a remote OS deployment, the system administrator must physically visit the system console or use a Dell remote access controller (RAC) to finish the PSSU. After the PSSU, the network security policy is unloaded and Windows Firewall services will be turned off to the default state.

TCP port 445 is blocked when Windows Firewall is first enabled. As a result, many of the Microsoft Management Console (MMC) snap-ins will fail when attempting to administer remote systems, as will the Find Users and Computers utility, resource kit utilities, and other utilities and third-party products that depend on the Server Message Block (SMB) protocol over TCP/IP. Examples of MMC snap-ins and utilities that depend on this TCP port include:

- Computer Management (compmgmt.msc)
- Device Manager (devmgmt.msc)
- Event Viewer (eventvwr.msc)
- Group Policy Results (gpresult.exe)
- Resultant Set of Policy (rsop.msc)
- Net services commands (net.exe)

Administrators who use Windows Terminal Server or Remote Desktop for Administration to remotely administer servers will also need to open TCP port 3389 unless they have configured Terminal Server to use an alternate port.

Toward successful upgrades to SP1

Unlike previous Microsoft OS service pack releases, Windows Server 2003 SP1 introduces major changes and features that can help significantly enhance the security of the OS. Carefully considering the deployment paths explored in this article and evaluating the application compatibility and server management issues identified will help administrators plan and execute the optimal route to smooth deployment in their organizations.

Min-John Lee is a software engineering consultant in the Server Operating Systems Engineering department in the Dell Product Group–Enterprise Software Development. Min-John has an M.S. in Electrical and Computer Engineering from Northwestern University.

Scott M. Callaway is a software engineer in the Server Operating Systems Engineering department in the Dell Product Group–Enterprise Software Development. Scott has a B.S. in Management from Stephen F. Austin State University.

Jeff Ferris is a manager in the Dell IT Engineering department. Jeff has a B.S. in Computer Information Systems from Southwest Missouri State University.

[14] For more information about how to use SCW, select Help and Support from the Start menu.

Planning Considerations for Multicore Processor Technology

The need to achieve higher performance without driving up power consumption and heat has become a critical concern for many IT organizations, given the density levels at which industry-standard servers are being deployed and the power and thermal constraints in today's data centers. Forthcoming multicore processor architectures will be designed to boost performance and minimize heat output by integrating two or more processor cores into a single processor socket. This article introduces the multicore concept and discusses key factors that IT organizations should consider when determining how best to take advantage of multicore technology.

BY JOHN FRUEHE

Related Categories:

Dell PowerEdge servers

Dell Precision workstations

Intel EM64T processors

Multiprocessor (MP)

Performance

Planning

Processors

Scalable enterprise

System architecture

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

Server density has grown dramatically over the past decade to keep pace with escalating performance requirements for enterprise applications. Ongoing progress in processor designs has enabled servers to continue delivering increased performance, which in turn helps fuel the powerful applications that support rapid business growth. However, increased performance incurs a corresponding increase in processor power consumption—and heat is a consequence of power use. As a result, administrators must determine not only how to supply large amounts of power to systems, but also how to contend with the large amounts of heat that these systems generate in the data center.

As more applications move from proprietary to standards-based systems, the performance demands on industry-standard servers are spiraling upward. Today, in place of midrange and large mainframe systems, tightly packed racks of stand-alone servers and blade servers can be clustered to handle the same types of business-critical application loads that once required large proprietary systems. Organizations are using databases such as Microsoft® SQL Server, Oracle® Database 10*g*, and MySQL to enhance business decision making along with enterprise-wide messaging applications such as Microsoft Exchange. Meanwhile, network infrastructure, Internet connectivity, and e-commerce are growing at tremendous rates. Altogether, the result is a steady increase in performance demands as user loads and processing loads grow, driving a steady increase in the density of systems in the data center, which is intensified by ever-faster processors—and in turn this can create power and cooling challenges for many IT organizations.

Current options to address power and cooling challenges

Historically, processor manufacturers have responded to the demand for more processing power primarily by delivering faster processor speeds. However, the challenge of managing power and cooling requirements for today's powerful processors has prompted a reevaluation of this approach to processor design. With heat rising incrementally faster than the rate at which signals move through the processor, known as clock speed, an increase in performance can create an even larger increase in heat.

IT organizations must therefore find ways to enhance the performance of databases, messaging applications, and other enterprise systems while contending with a corresponding increase in system power consumption and heat. Although faster processors are one way to improve server performance, other approaches can help boost performance without increasing clock speed and incurring an attendant increase in power consumption and heat. In fact, excellent overall processing performance may be achieved by *reducing* clock speed while increasing the number of processing units—and the consequent reduction in clock speed can lead to lower heat output and greater efficiency. For example, by moving from a single high-speed core, which generates a corresponding increase in heat, to multiple slower cores, which produce a corresponding reduction in heat, enterprises can potentially improve application performance while reducing their thermal output.

Balancing performance across each platform. The first step is to optimize performance across all platform elements. Designing, integrating, and building complete platforms that balance computing capabilities across processor, chip set, memory, and I/O components can significantly improve overall application performance and responsiveness. By integrating flexible technologies and balancing performance across all platform components, administrators can help provide the headroom required to support business growth (such as increases in users, transactions, and data) without having to upgrade the entire server. This approach can help the systems in place today support increased business demands, enhancing scalability for future growth. At the same time, this strategy can help extend the life of existing data center components by enabling administrators to optimize the performance of repurposed platforms when next-generation applications are deployed.

Harnessing multithreading technology. The second step is to improve the efficiency of computer platforms by harnessing the power of multithreading. Industry-standard servers with multiple processors have been available for many years, and the overwhelming majority of networked applications can take advantage of the additional processors, multiple software threads, and multitasked computing environments. These capabilities have enabled organizations to scale networked applications for greater performance. The next logical step for multiprocessing advancements is expected to come in the form of multiple logical processing units, or *processor cores,* within a single chip. Multicore processors—coupled with advances in memory, I/O, and storage—can be designed to deliver a balanced platform that enables the requisite performance and scalability for future growth.

Optimizing software applications. The third step, software optimization, can be an efficient way to enable incremental performance gains without increasing power consumption and heat. Many of today's leading software tools, along with Intel® compilers, can enable significant performance improvements over applications that have not been compiled or tuned using such optimization tools.[1] Actual performance gains will depend on the specific system configuration and application environment. To get the most performance from existing data center components, administrators must not overlook potential gains from optimizing software applications during the infrastructure planning processes.

Scalability potential of multicore processors

Processors plug into the system board through a socket. Current technology allows for one processor socket to provide access to one logical core. But this approach is expected to change, enabling one processor socket to provide access to two, four, or more processor cores. Future processors will be designed to allow multiple processor cores to be contained inside a single processor module. For example, a tightly coupled set of dual processor cores could be designed to compute independently of each other—allowing applications to interact with the processor cores as two separate processors even though they share a single socket. This design would allow the OS to "thread" the application across the multiple processor cores and could help improve processing efficiency.

A multicore structure would also include cache modules. These modules could either be shared or independent. Actual implementations of multicore processors would vary depending on manufacturer and product development over time. Variations may include shared or independent cache modules, bus implementations, and additional threading capabilities such as Intel Hyper-Threading (HT) Technology.

A multicore arrangement that provides two or more low-clock-speed cores could be designed to provide excellent performance while minimizing power consumption and delivering lower heat output than configurations that rely on a single high-clock-speed core. The following example shows how multicore technology could manifest in a standard server configuration and how multiple

[1] For example, in January 2005 Intel conducted benchmark tests showing that 64-bit Intel Xeon processor technology with Intel Hyper-Threading Technology enabled can provide up to 33 percent improvement in application and server performance compared to the same configuration with Hyper-Threading Technology disabled. For more information about Hyper-Threading Technology performance tests, visit www.intel.com/performance/server/xeon/ht_perf.htm. Please note that Intel performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Administrators should consult other sources of information to evaluate the performance of specific systems or components. For more information about performance tests and the performance of Intel products, visit www.intel.com/performance/resources/benchmark_limitations.htm or call (U.S.) 1-800-628-8686 or 1-916-356-3104.

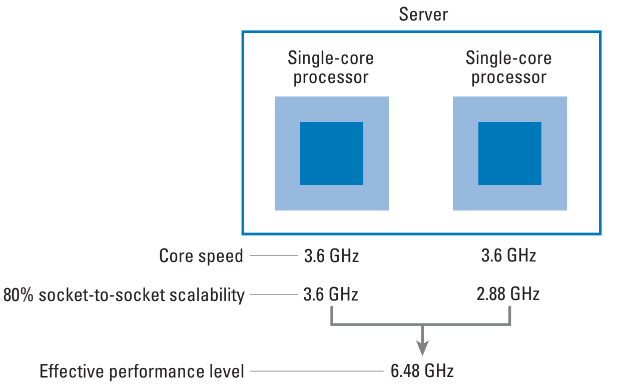


Figure 1. Sample core speed and anticipated total relative power in a system using two single-core processors

low-clock-speed cores could deliver greater performance than a single high-clock-speed core for networked applications.

This example uses some simple math and basic assumptions about the scaling of multiple processors and is included for demonstration purposes only. Until multicore processors are available, scaling and performance can only be estimated based on technical models. The example described in this article shows one possible method of addressing relative performance levels as the industry begins to move from platforms based on single-core processors to platforms based on multicore processors. Other methods are possible, and actual processor performance and processor scalability are tied to a variety of platform variables, including the specific configuration and application environment. Several factors can potentially affect the internal scalability of multiple cores, such as the system compiler as well as architectural considerations including memory, I/O, frontside bus (FSB), chip set, and so on.

For instance, enterprises can buy a dual-processor server today to run Microsoft Exchange and provide e-mail, calendaring, and messaging functions. Dual-processor servers are designed to deliver excellent price/performance for messaging applications. A typical configuration might use dual 3.6 GHz 64-bit Intel Xeon™ processors supporting HT Technology. In the future, organizations might deploy the same application on a similar server that instead uses a pair of dual-core processors at a clock speed lower than 3.6 GHz. The four cores in this example configuration might each run at 2.8 GHz. The following simple example can help explain the relative performance of a low-clock-speed, dual-core processor versus a high-clock-speed, dual-processor counterpart.

Dual-processor systems available today offer a scalability of roughly 80 percent for the second processor, depending on the OS, application, compiler, and other factors.[2] That means the first processor may deliver 100 percent of its processing power, but the second processor typically suffers some overhead from multiprocessing activities. As a result, the two processors do not scale linearly—that is, a dual-processor system does not achieve a 200 percent performance increase over a single-processor system, but instead provides approximately 180 percent of the performance that a single-processor system provides. In this article, the single-core scalability factor is referred to as external, or *socket-to-socket,* scalability. When comparing two single-core processors in two individual sockets, the dual 3.6 GHz processors would result in an effective performance level of approximately 6.48 GHz (see Figure 1).

For multicore processors, administrators must take into account not only socket-to-socket scalability but also internal, or *core-to-core,* scalability—the scalability between multiple cores that reside within the same processor module. In this example, core-to-core scalability is estimated at 70 percent, meaning that the second core delivers 70 percent of its processing power. Thus, in the example system using 2.8 GHz dual-core processors, each dual-core processor would behave more like a 4.76 GHz processor when the performance of the two cores—2.8 GHz plus 1.96 GHz—is combined.

For demonstration purposes, this example assumes that, in a server that combines two such dual-core processors within the same system architecture, the socket-to-socket scalability of the two dual-core processors would be similar to that in a server containing two single-core processors—80 percent scalability. This would lead to an effective performance level of 8.57 GHz (see Figure 2).

To continue the example comparison by postulating that socket-to-socket scalability would be the same for these two architectures, a

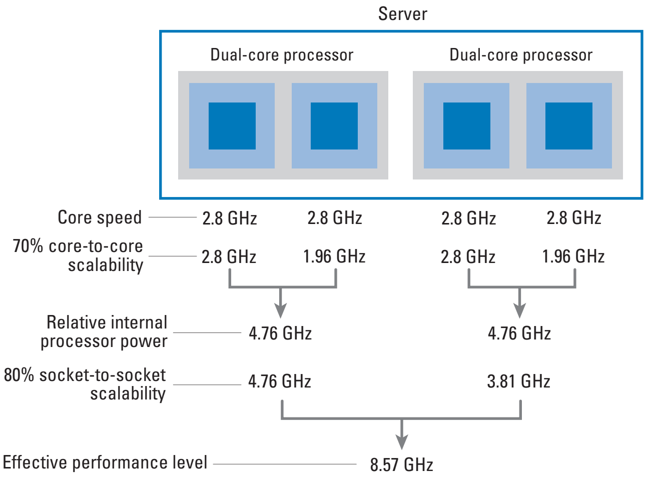


Figure 2. Sample core speed and anticipated total relative power in a system using two dual-core processors

[2] While 80 percent scalability for the second processor is a representative approximation for the example set forth in this article, even higher scalability has been achieved. For example, in January 2005 Intel conducted benchmark tests of 64-bit Intel Xeon processor scaling based on a two-processor configuration versus an otherwise comparable one-processor configuration. For more information, visit www.intel.com/performance/server/xeon/scaling.htm.

UNDERSTANDING HYPER-THREADING TECHNOLOGY

Today's 64-bit Intel Xeon, Pentium® 4, and Celeron® processors include HT Technology, which enables the processor to execute multiple threads of an application simultaneously. Multithreaded applications perceive a single physical processor as two separate, logical processors and will execute threads independently on each logical processor to help speed overall processing execution. Recent benchmark tests by Intel of 64-bit Intel Xeon processor–based platforms have shown a performance gain of up to 33 percent by enabling HT Technology on applications that are HT Technology–aware as compared to running the same applications with HT Technology disabled.*

Today, individual Intel NetBurst® microprocessors appear to the OS as two logical processors. On a dual-processor system supporting HT Technology, the application perceives four processor threads (two physical processors and two logical processors). Equipped with multicore processors, that same dual-socket system could have a total of four processor cores. Through the effective use of HT Technology, those four processor cores could appear to the application as eight total processors.

By leveraging HT Technology, a properly compiled application can achieve performance increases because of the improved utilization of the existing system processors, compared to the same application running with HT Technology disabled. Most multiprocessor-aware applications can take advantage of HT Technology, and applications that have been specifically designed for HT Technology have the potential to achieve a significant performance increase.

By combining multicore processors with HT Technology, Intel aims to provide greater scalability and better utilization of processing cycles within the server than is possible using single-core processors with HT Technology. The addition of HT Technology to multicore processor architecture could present an excellent opportunity to help improve the utilization and scalability of future processor subsystems.

*For more information about Intel HT Technology performance tests, visit www.intel.com/performance/server/xeon/ht_perf.htm. Please note that Intel performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Administrators should consult other sources of information to evaluate the performance of specific systems or components. For more information about performance tests and the performance of Intel products, visit www.intel.com/performance/resources/benchmark_limitations.htm or call (U.S.) 1-800-628-8686 or 1-916-356-3104.

multicore architecture could enable greater performance than a single-core processor architecture, even if the processor cores in the multicore architecture are running at a lower clock speed than the processor cores in the single-core architecture. In this way, a multicore architecture has the potential to deliver higher performance than a single-core architecture for enterprise applications.

Power and cooling advantages of multicore processors

Although the preceding example explains the scalability potential of multicore processors, scalability is only part of the challenge for IT organizations. High server density in the data center can create significant power consumption and cooling requirements. A multicore architecture can help alleviate the environmental challenges created by high-clock-speed, single-core processors.

Heat is a function of several factors, two of which are processor density and clock speed. Other drivers include cache size and the size of the core itself. In traditional architectures, heat generated by each new generation of processors has increased at a greater rate than clock speed.

In contrast, by using a shared cache (rather than separate dedicated caches for each processor core) and low-clock-speed processors, multicore processors may help administrators minimize heat while maintaining high overall performance. This capability may help make future multicore processors attractive for IT deployments in which density is a key factor, such as high-performance computing (HPC) clusters, Web farms, and large clustered applications. Environments in which 1U servers or blade servers are being deployed today could be enhanced by potential power savings and potential heat reductions from multicore processors.

Currently, technologies such as demand-based switching (DBS) are beginning to enter the mainstream, helping organizations reduce the utility power and cooling costs of computing. DBS allows a processor to reduce power consumption (by lowering frequency and voltage) during periods of low computing demand. In addition to potential performance advances, multicore designs also hold great promise for reducing the power and cooling costs of computing, given DBS technology. DBS is available in single-core processors today, and its inclusion in multicore processors may add capabilities for managing power consumption and, ultimately, heat output. This potential utility cost savings could help accelerate the movement from proprietary platforms to energy-efficient industry-standard platforms.

Significance of sockets in a multicore architecture

As they become available, multicore processors will require IT organizations to consider system architectures for industry-standard servers from a different perspective. For example, administrators currently segregate applications into single-processor, dual-processor, and quad-processor classes. However, multicore processors will call for a new mind-set that considers processor cores as well as sockets.

Single-threaded applications that perform best today in a single-processor environment will likely continue to be deployed on single-processor, single-core system architectures. For single-threaded applications, which cannot make use of multiple processors in a system, moving to a multiprocessor, multicore architecture may not necessarily enhance performance. Most of today's leading operating systems, including Microsoft Windows Server System™ and Linux® variants, are multithreaded, so multiple single-threaded applications can run on a multicore architecture even though they are not inherently multithreaded. However, for multithreaded applications that are currently deployed on single-processor architectures because of cost constraints, moving to a single-processor, dual-core architecture has the potential to offer performance benefits while helping to keep costs low.

For the bulk of the network infrastructure and business applications that organizations run today on dual-processor servers, the computing landscape is expected to change over time. However, while it may initially seem that applications running on a dual-processor, single-core system architecture can migrate to a single-processor, dual-core system architecture as a cost-saving initiative, this is not necessarily the case. To maintain equivalent performance or achieve a greater level of performance, the dual-processor applications of today will likely have to migrate to dual-socket, dual-core systems. As postulated in the Figure 1 example, a system architecture consisting of four processor cores in two sockets can be designed to deliver superior performance relative to a dual-socket, single-core system architecture, while also delivering potential power and cooling savings to the data center. The potential to gradually migrate a large number of older dual-socket, single-core servers to energy-efficient dual-socket, multicore systems could enable significant savings in power and cooling costs over time. Because higher-powered, dual-socket systems typically run applications that are more mission-critical than those running on less-powerful, single-processor systems, organizations may continue to expect more availability, scalability, and performance features to be designed for dual-socket systems relative to single-socket systems—just as they do today.

For applications running today on high-performing quad-processor systems, a transition to multicore technology is not necessarily an opportunity to move from four-socket, four-core systems to dual-socket, four-core systems. Rather, the architectural change suggests that today's four-processor applications may migrate to four-socket systems with eight or potentially more processor cores—helping to extend the range of cost-effective, industry-standard alternatives to large, proprietary symmetric multiprocessing (SMP) systems. Because quad-processor systems tend to run more mission-critical applications in the data center as compared to dual-processor systems and single-processor systems, administrators can expect quad-processor platforms to be designed with the widest range of performance, availability, and scalability features across Dell™ PowerEdge™ server offerings.

When comparing relative processing performance of one generation of servers to the next, a direct comparison should not focus on the number of processor cores but rather on the number of sockets. However, the most effective comparison is ultimately not one of processors or sockets alone, but a thorough comparison of the entire platform—including scalability, availability, memory, I/O, and other features. By considering the entire platform and all the computing components that participate in it, organizations can best match a platform to their specific application and business needs.

Evolution of software toward multicore technology

Multicore processing continues to exert a significant impact on software evolution. Before the advent of multicore processor technology, both SMP systems and HT Technology motivated many OS and application vendors to design software that could take advantage of multithreading capabilities. As multicore processor–based systems enter the mainstream and evolve, it is likely that OS and application vendors will optimize their offerings for multicore architectures, resulting in potential performance increases over time through enhanced software efficiency.

> Although faster processors are one way to improve server performance, other approaches can help boost performance without increasing clock speed and incurring an attendant increase in power consumption and heat.

Most application vendors will likely continue to develop on industry-standard processor platforms, considering the power, flexibility, and huge installed base of these systems. Currently, 64-bit Intel Xeon processors have the capability to run both 32-bit applications and 64-bit applications through the use of Intel Extended Memory 64 Technology (EM64T). The industry is gradually making the transition from a 32-bit standard to a 64-bit standard, and similarly, software can be expected to make the transition to take advantage of multicore processors over time.

Applications that are designed for a multiprocessor or multithreaded environment can currently take advantage of multicore processor architectures. However, as software becomes optimized for multicore processors, organizations can expect to see overall application performance enhancements deriving from software innovations that take advantage of multicore-processor–based system architecture instead of increased clock speed.

In addition, compilers and application development tools will likely become available to optimize software code for multicore processors, enabling long-term optimization and enhanced efficiency for multicore processors—which also may help realize performance improvements through highly tuned software design rather than a brute-force increase in clock speed. Intel is working toward introducing software tools and compilers to help optimize threading performance for both single-core and multicore architectures. Organizations that begin to optimize their software today for multicore system architecture may gain significant business advantages as these systems become mainstream over the next few years. For instance, today's dual Intel Xeon processor–based system with HT Technology can support four concurrent threads (two per processor). With the advent of dual-core Intel Xeon processors with HT Technology, these four threads would double to eight. An OS would then have eight concurrent threads to distribute and manage workloads, leading to potential performance increases in processor utilization and processing efficiency.

Licensing considerations

Another key area to consider in planning for a migration to multicore processors is the way in which software vendors license their applications. Many enterprise application vendors license their applications based on the number of processors, not the number of users. This could mean that, although a dual-socket, dual-core server may offer enhanced performance when compared to a dual-socket, single-core server, the licensing cost could potentially double because the application would identify four processors instead of two. The resulting increase in licensing costs could negate the potential performance improvement of using multicore processor–based systems. Because the scalability of multicore processors is not linear—that is, adding a second core does not result in a 100 percent increase in performance—a doubling of licensing costs would result in lower overall price/performance.

For that reason, software licensing should be considered a key factor when organizations assess which applications to migrate to systems using multicore processors. For example, enterprise software licensing costs can be significantly higher than the cost of the server on which the application is running. This can be especially true for industry-standard servers that deliver excellent performance at a low price point as compared to proprietary servers. Some application vendors have adopted a policy of licensing based on the socket count instead of the number of cores, while others have not yet taken a stance on this matter. Until the industry gains more clarity around this software licensing issue, organizations must factor software licensing costs into the overall platform cost when evaluating multicore technology transitions.

Dell multicore processor plans

Dell plans to begin integrating multicore processor designs into Dell PowerEdge and Dell PowerEdge SC servers during the next 12 to 18 months as multicore processors become available. Through Dell's close relationship with Intel, Dell intends to deliver solutions built on Intel Xeon, Pentium 4, and Itanium® multicore processors. Dell is working closely with Intel to ensure that the next generations of PowerEdge and PowerEdge SC servers are designed to meet both the performance and scalability needs of enterprises in relation to multicore processor architectures.

Shift in focus toward multicore technology

Multicore processors most likely represent the future direction of server architecture, which is expected to enhance application performance and platform power with thermal efficiency. By combining multiple logical processing units within a single processor package as described in this article, multicore processor architectures have the potential to provide superior performance and scalability without a corresponding increase in power consumption and heat, as would be the case by simply increasing the clock speed of existing single-core processor designs.

In readiness for this impending change in processor architecture, system vendors like Dell will begin to address system design in a different manner when determining system performance and scalability. What was once a focus on processor count will become a focus on socket count as the shift occurs in the number of processor cores per socket. However, before adopting this emerging system architecture, IT organizations need to carefully evaluate the software ramifications of migrating applications to multicore processor technology. This consideration can enable enterprises to benefit from the higher performance and lower power consumption expected of multicore processor architecture as compared to single-core processor architecture, while helping ensure that multicore processor platforms are licensed appropriately to control acquisition costs.

John Fruehe is a marketing strategist for the Dell Enterprise Product Group. He has worked at Dell for nine years. Prior to that, John was at Compaq and Zenith Data Systems. John has a B.S. in Economics from Illinois State University and has been in the technology field for 14 years.

FOR MORE INFORMATION

Multicore processor architecture:
www.intel.com/cd/ids/developer/asmo-na/eng/201969.htm?page=6

Dual-core and HT Technology:
www.intel.com/cd/ids/developer/asmo-na/eng/technologies/threading/199701.htm

Achieving High Availability

with the Turnkey Dell PowerEdge Cluster FE500W-IA64

To support business-critical applications with a high level of availability, scalability, and reliability, IT administrators can cluster 64-bit Dell™ PowerEdge™ 7250 servers using the 64-bit version of the Microsoft® Windows Server™ 2003, Enterprise Edition, operating system. This article describes the turnkey Dell PowerEdge Cluster FE500W-IA64 cluster configuration, including Dell PowerEdge 7250 servers and Dell/EMC storage components.

BY BRYANT VO, NAM NGUYEN, AND DAT NGUYEN

Related Categories:

Clustering

Dell PowerEdge servers

Dell PowerVault storage

Dell/EMC storage

Fibre Channel

High availability (HA)

Intel IA-64 (Itanium) processors

Microsoft Windows

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

The Dell PowerEdge Cluster FE500W-IA64 is the first Microsoft Cluster Server (MSCS) implementation developed by Dell for 64-bit applications. This 64-bit cluster incorporates Microsoft Windows Server 2003, Enterprise Edition for 64-Bit Itanium-based Systems, and 64-bit Dell PowerEdge 7250 servers that are each equipped with up to four Intel® Itanium® 2 processors. Dell PowerEdge 7250 servers offer wide system buses, enhanced memory addressing, and enhanced parallelism, making them well suited for compute-intensive native 64-bit applications that require high levels of availability and scalability.

Overview of high-availability clusters

High-availability clusters—often referred to as failover clusters—use specific hardware and software designed to link multiple systems so that these systems can function together as a single virtual system. If a hardware or software fault causes one node in a cluster to fail, resources running on the failed node are moved to one or more remaining nodes in the cluster, a process known as failover. Typically, users connected to the failed node experience only a momentary delay while cluster software reestablishes the connection with another node in the cluster and restarts applications on that node.

MSCS is the failover software component available in specific Windows® operating systems. MSCS is designed to provide high availability to applications and services such as databases, messaging, file servers, print servers, and basic Web servers. MSCS works well with Dell hardware and software components—including the Dell PowerEdge Cluster FE500W-IA64 configuration—to provide high availability and scalability for business-critical applications.

Dell PowerEdge Cluster FE500W-IA64 components

The Dell PowerEdge Cluster FE500W-IA64 is designed for large organizations running business-critical database applications that require a high level of availability, scalability, and reliability. The cluster helps achieve these criteria by incorporating the following components: Microsoft Windows Server 2003, Enterprise Edition for 64-Bit Itanium-based Systems; up to eight Dell PowerEdge 7250 servers; Dell/EMC CX300, CX500, and CX700 Fibre Channel storage arrays; and Dell/EMC storage management software.

The Dell PowerEdge Cluster FE500W-IA64 also supports the following cluster applications: Microsoft SQL Server 2000 Enterprise Edition (64-bit); Oracle9*i*™ Database; Oracle Database 10*g* with Oracle Fail Safe release 3.3.2 for Windows Server 2003 (64-bit); and VERITAS NetBackup.

Servers and operating system. Equipped with Intel Itanium 2 processors, the Dell PowerEdge 7250 can be an excellent server for memory-intensive and processor-intensive applications. The server can also be optimal for database applications requiring a high level of compute-parallelism, scalability, and reliability. Designed with Explicitly Parallel Instruction Computing (EPIC) architecture, 6.4 Gbps system bus bandwidth, up to 9 MB of level 3 (L3) cache, and 64-bit addressing, the server's Intel Itanium 2 processors can greatly enhance performance and scalability for processing large amounts of data per clock cycle, addressing memory, and running numeric calculations. In addition, the Dell PowerEdge 7250 server supports industry-standard technology, which can make it a cost-effective alternative to proprietary RISC/UNIX–based systems—thereby helping to provide fast return on investment and low total cost of ownership.

> High-availability clusters—often referred to as failover clusters—use specific hardware and software designed to link multiple systems so that these systems can function together as a single virtual system.

Each Dell PowerEdge 7250 in the Dell PowerEdge Cluster FE500W-IA64 includes the following: up to four Intel Itanium 2 processors; up to 64 GB of double data rate (DDR) error-correcting code (ECC) SDRAM; Windows Server 2003, Enterprise Edition for 64-Bit Itanium-based Systems; redundant hot-plug hard drives, power supplies, and cooling fans; modular 4U chassis design; redundant QLogic QLA2340 host bus adapters (HBAs); and redundant paths to the storage systems.

Storage systems. The Dell/EMC CX-series Fibre Channel storage arrays supported in the cluster include the CX300, CX500, and CX700 models. A Dell PowerEdge Cluster FE500W-IA64 is designed to support up to four storage arrays in a cluster.

Each array in a Dell PowerEdge Cluster FE500W-IA64 includes the following features: 2 Gbps Fibre Channel technology; global hot spares; redundant storage processors; redundant back-end loops with dual-port Fibre Channel drives; redundant Fibre Channel fabrics; nondisruptive upgrade (NDU) capabilities; consolidated backup implementations; and hot-swappable components, including drives, power supplies, cooling fans, link control cards, and storage processors.

Figure 1 lists specifications for the Dell PowerEdge Cluster FE500W-IA64–supported Dell/EMC storage arrays.

Storage management software. The Dell PowerEdge Cluster FE500W-IA64 incorporates the following Dell/EMC storage management software to handle communications between the nodes and the storage systems in a storage area network (SAN) environment:

- **EMC Access Logix:** Enables multiple-node connectivity and storage consolidation by allowing multiple nodes and servers to share a Dell/EMC storage system. Access Logix™ software is designed to restrict server access to specific volumes on a shared storage system to protect data from unauthorized access.
- **EMC PowerPath:** Provides multiple-path I/O capabilities, automatic load balancing, and path failover. PowerPath® software helps detect I/O path failures and is designed to automatically reroute I/Os through an alternate path when failures occur.
- **EMC MirrorView:** Offers remote synchronous or asynchronous mirroring to help provide business continuity. MirrorView™ software enables highly available data mirroring over distances by supporting both IP network and Fibre Channel connections.
- **EMC SnapView:** Creates multiple, point-in-time copies of production data for fast data backup and recovery. SnapView™ software captures images of a logical storage unit (LUN); these images can be used to share LUNs across nodes without affecting the contents of the source LUN.
- **EMC SAN Copy:** Provides data mobility between different storage systems over a high-speed SAN infrastructure or wide area network (WAN). SAN Copy™ software is designed to move data between storage systems without using host processor cycles or LAN bandwidth. It can be used in conjunction with MirrorView or SnapView.

Dell PowerEdge Cluster FE500W-IA64 storage configuration options

The Dell PowerEdge Cluster FE500W-IA64 supports both direct attach and SAN-attached environments using the following components: up to eight Dell PowerEdge 7250 nodes in a cluster;

Feature	CX300	CX500	CX700
Throughput	50,000 I/Os per second (IOPS)	120,000 IOPS	200,000 IOPS
Bandwidth	680 Mbps	760 Mbps	1,520 Mbps
Number of disks	60	120	240
Maximum capacity	9 TB	18 TB	35 TB
System cache	2 GB	4 GB	8 GB
Maximum LUNs	512	1,024	2,048
Maximum number of high-availability nodes per array	64	128	256

Figure 1. Dell/EMC CX300, CX500, and CX700 array specifications

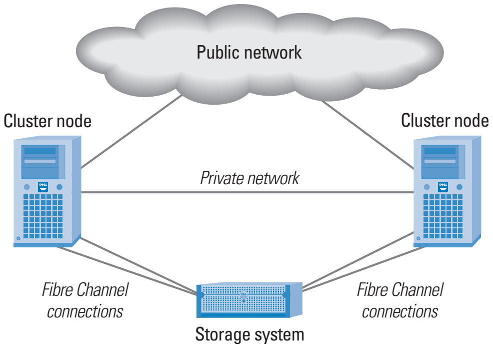


Figure 2. Direct attach cluster configuration

Microsoft Windows Server 2003, Enterprise Edition for 64-Bit Itanium-based Systems; platform-supported Gigabit[1] Ethernet network interface cards (NICs); QLogic QLA2340 HBAs; Brocade SilkWorm 3200 and 3800 Fibre Channel switches; Dell/EMC CX300, CX500, and CX700 storage arrays; and Dell/EMC SAN management and path failover software.

Direct attach cluster. In a direct attach cluster configuration, each node of the cluster is directly attached to a single storage array. In this configuration, the Fibre Channel HBAs in the cluster nodes are cabled directly to the storage processors on the Dell/EMC array. Figure 2 shows a basic direct attach, single-cluster configuration with two nodes.

SAN-attached cluster. In a SAN-attached cluster, up to eight nodes are attached to up to four storage arrays through a SAN using a redundant switch fabric. This configuration is designed to enhance the functionality of SAN-attached clusters, making them superior to direct attach clusters in configuration flexibility, expandability, and performance. Figure 3 shows a SAN-attached cluster.

Fully integrated clustering product

The Dell PowerEdge Cluster FE500W-IA64 offers an integrated and validated high-availability clustering implementation that helps provide IT organizations with an excellent approach for clustering 64-bit applications. This 64-bit Fibre Channel cluster is designed to deliver a high level of availability and scalability that can benefit 64-bit workloads, including large databases and other business-critical applications that require continuous availability.

Bryant Vo is a systems engineer and consultant in the High-Availability Cluster Development Group at Dell. His current projects include MSCS clustering and SANs. He has a B.S. in Computer Science from the University of Houston.

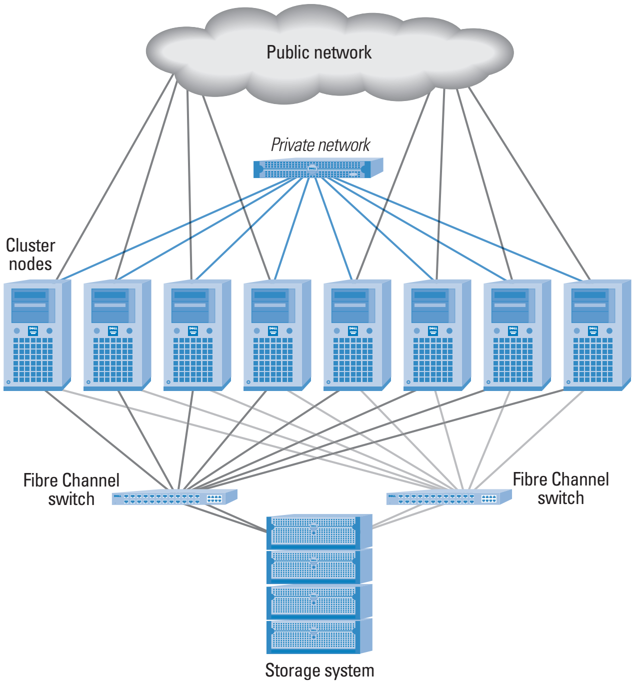


Figure 3. SAN-attached cluster configuration

Nam Nguyen is a senior consultant in the High-Availability Cluster Development Group at Dell, and the lead engineer for Dell Fibre Channel PowerEdge Cluster products. His current interests include business continuity, clustering, and storage technologies. He has a B.S. and an M.S. in Electrical Engineering from The University of Texas at Austin.

Dat Nguyen is a systems engineer in the High-Availability Cluster Development Group at Dell. His responsibilities include developing SAN-based high-availability clustering products that comprise Dell servers and Dell/EMC Fibre Channel storage systems. Dat has a B.S. in Electrical Engineering from the University of Houston.

FOR MORE INFORMATION

Dell high availability clustering:
www.dell.com/ha

Dell cluster configuration support matrices:
www1.us.dell.com/content/topics/global.aspx/solutions/en/clustering_ha?c=us&l=en&s=gen&~tab=3

Intel Itanium 2 processor–based server platforms:
www.intel.com/business/bss/products/server/itanium2/index.htm

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

Implementing Oracle Database 10*g*

Maximum Availability Architecture

on Dell PowerEdge Servers and Dell/EMC Storage over Wide Area Networks

The high cost of downtime has prompted many organizations to view business continuity and high availability as critical IT concerns. This article explores how components of Oracle® Maximum Availability Architecture, including Oracle Real Application Clusters and Oracle Data Guard, can be implemented on Dell™ PowerEdge™ servers and Dell/EMC storage to help build the foundation of a scalable, end-to-end architecture.

BY MICHAEL SMITH, PAUL RAD, AND ASHISH RAY

Related Categories:

Clustering

Database

Dell PowerEdge servers

Dell/EMC storage

Disaster recovery

High availability (HA)

Oracle

Storage

Wide area network (WAN)

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

If a critical application, server, or storage subsystem fails—or an unforeseen disaster strikes—the ensuing downtime can place an enterprise in jeopardy. Even carefully planned downtime can affect user productivity when mission-critical applications must be taken offline. Whether planned or unplanned, downtime can translate into lost business opportunities and increased costs.

To help address business continuity requirements, engineering teams from Dell and Oracle conducted a joint project to demonstrate best practices for building a high-availability architecture that enables enterprises to minimize downtime for business applications and build resiliency within their IT infrastructure. This architecture used the Real Application Clusters (RAC) and Data Guard features of Oracle Database 10*g* in conjunction with Dell PowerEdge servers and Dell/EMC storage to help provide high availability, scalability, and disaster protection.

This article explores the high-availability features enabled by Oracle Database 10g, Dell PowerEdge servers, and Dell/EMC storage. In addition, the article describes best practices used in the Dell and Oracle joint project.

Integrated high-availability features

One challenge in designing a high-availability IT infrastructure is examining and addressing all possible causes of downtime. Downtime can be classified into two primary categories: unplanned and planned. IT organizations should consider potential causes of both unplanned and planned downtime when designing a fault-tolerant and resilient IT infrastructure. Unplanned downtime primarily results from system failures or data failures. Planned downtime is typically caused by system changes or data changes that must be applied to the production system.

As shown in Figure 1, Oracle Database 10*g*, Dell PowerEdge servers, and Dell/EMC storage offer an integrated set of high-availability features designed to help organizations minimize the various kinds of downtime that can affect their businesses.[1]

[1] For an overview of the high-availability features of Oracle Database 10*g*, visit www.oracle.com/technology/deploy/availability. For Oracle high-availability architecture and best-practices documentation, visit download-west.oracle.com/docs/cd/B14117_01/server.101/b10726/toc.htm and www.dell.com/oracle.

Minimizing unplanned downtime

To help protect against server failures, Oracle RAC allows multiple Dell PowerEdge servers to access a single Oracle database in a clustered environment. This approach can offer the benefit of scalability without requiring changes in application code.

To help protect against downtime caused by various problems—including storage failure, human error, data corruption, and site disruption—Oracle Database 10*g* offers a suite of features. Among them, Automatic Storage Management (ASM) offers file system and volume manager capabilities integrated with Oracle Database 10*g* as well as native mirroring of database files for enhanced protection.[2] To help protect against human error, Oracle Database 10*g* offers the Flashback suite of features. For example, Flashback Database and Flashback Table enable administrators to easily "rewind" the state of the database or database objects to a known, safe point in time. This approach can help undo the effects of human error without incurring much downtime.

To help protect data from various media failures, Oracle offers Recovery Manager (RMAN), which is designed to provide comprehensive backup, restore, and recovery capabilities for the Oracle database. With RMAN, Oracle database backups can be performed online to avoid incurring expensive downtime. Furthermore, Oracle Database 10*g* offers the Flash Recovery Area, which is a unified disk-based storage location for all recovery-related files and activities in an Oracle database. RMAN and Flash Recovery Area are designed to work together to allow an enhanced, automated approach to disk-based backup and recovery. This approach enables administrators to perform fast backups and restores of the Oracle database to help shrink maintenance windows.

Finally, Oracle offers Data Guard to help protect against site disruptions or storage system failures that could result from localized or regional disasters such as fires, earthquakes, hurricanes, and malicious acts. Data Guard enables multiple *standby* databases to be connected to the production, or *primary,* database over a network, keeping the standby databases transactionally consistent with the primary database. If an unforeseen disaster occurs at the primary data center, Data Guard allows the production role to be switched easily to a chosen standby database, and Data Guard can be configured to avoid data loss during this process. Data Guard also allows the standby databases to be used for activities such as reporting and backup.

Dell/EMC storage arrays such as the Dell/EMC CX300, CX500, and CX700 are designed to help improve availability and data integrity by removing single points of failure.[3]

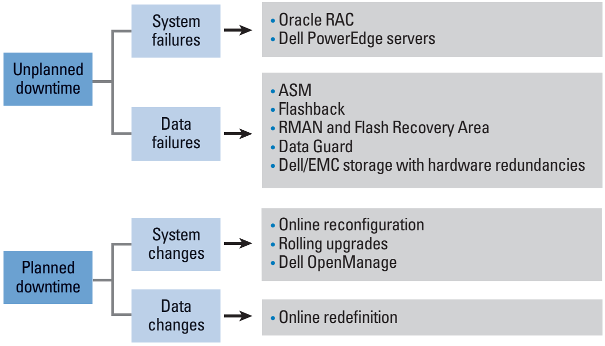


Figure 1. Dell and Oracle integrated high-availability features

Minimizing planned downtime

Planned downtime, which includes activities such as routine maintenance and new deployments, can be just as disruptive to operations as unplanned downtime—especially in enterprises that support users across several time zones. Besides helping to minimize unplanned downtime, Oracle Database 10*g* and Dell PowerEdge servers offer a suite of capabilities that can minimize or eliminate planned downtime.

For example, the Oracle database dynamically accommodates changes to hardware configurations such as adding and removing nodes in a RAC cluster, dynamically growing and shrinking shared memory allocation, and adding and removing online database disks without disturbing database activities. Moreover, the rolling upgrades capability uses Data Guard SQL Apply to permit upgrades of database patch sets or major releases[4] in rotation, which also helps minimize application downtime.

Dell OpenManage™ infrastructure provides administrators with comprehensive, one-to-one systems management capabilities for Dell PowerEdge servers to help eliminate planned hardware downtime and avoid hardware failures that could lead to unplanned downtime.[5]

Finally, to guard against problems caused by data changes, the online redefinition capability enables Oracle Database 10*g* to support many data maintenance operations without disrupting database operations or preventing users from updating or accessing data. For example, the online redefinition capability allows administrators to redefine database tables—including changing table types; adding, dropping, or renaming columns; and changing storage parameters—without interrupting end-user activities such as viewing and updating the underlying data.

[2] For more information about best practices for ASM on Dell/EMC storage, see "Best Practices for Oracle Database 10*g* Automatic Storage Management on Dell/EMC Storage" by Paul Rad, Ramesh Rajagopalan, Tesfamariam Michael, and Jay Kozak in *Dell Power Solutions,* October 2004.

[3] For more information about Dell/EMC storage, visit www.dell.com/emc.

[4] This capability enables upgrades for database patch sets or major releases from Oracle Database 10*g* Release 1 and later.

[5] For more information about Dell OpenManage, visit www.dell.com/openmanage.

Best-practices guidelines for configuring Oracle MAA on Dell servers

The goal of the Dell and Oracle project was to provide best-practices guidelines for configuring Oracle Maximum Availability Architecture (MAA) using Oracle Database 10*g* on Dell servers and Dell/EMC storage. The objective of MAA—Oracle's blueprint for Oracle high-availability technologies—is to remove the complexity of designing an optimal high-availability architecture and to help maximize systems availability.[6]

Two fundamental technologies that enable MAA are RAC[7] and Data Guard,[8] which help the system architecture provide end-to-end support for high availability and disaster protection. RAC is designed to provide IT organizations with near-instantaneous server failover when server failures occur, while Data Guard—especially if deployed over a wide area network (WAN)—helps protect production data from disasters that could otherwise severely affect data center operations. The following sections provide further details on the joint Dell and Oracle project.

System and network considerations

Conducted in the fourth quarter of 2004 by Dell and Oracle engineers, the joint Dell and Oracle project involved setting up three sites with identical hardware and software configurations: two sites in the Dell Engineering Lab in Austin, Texas, and one site in the Oracle System Technology Data Center in Redwood Shores, California (see Figure 2).

In the Dell and Oracle implementation configured for this study, each site consisted of redundant hardware and software components designed to ensure that all requests were serviced, even if a failure occurred. The primary site was located in Austin, with the primary database hosted on a two-node Dell cluster comprising Dell PowerEdge 2650 servers. The two servers were configured with Oracle Database 10*g* (10.1.0.2), including the RAC option, and a Dell/EMC CX500 Fibre Channel storage array.

Dell and Oracle engineers configured the implementation described in this article for demonstration purposes. However, many different implementations are possible using Dell and Oracle supported and validated configurations;[9] administrators should implement the configuration that best suits their organization's specific needs.

The primary database in the Dell and Oracle implementation described in this study was configured with two standby databases: a logical standby database and a physical standby database. The LAN-attached logical standby database was hosted in Austin and maintained using Data Guard to enable synchronous transport. The WAN-attached physical standby database was hosted in Redwood Shores and maintained using Data Guard to enable asynchronous transport. To allow the standby sites to process application workloads at the same performance level as the primary site during a role transition, both standby databases were configured identically to the database at the primary site.

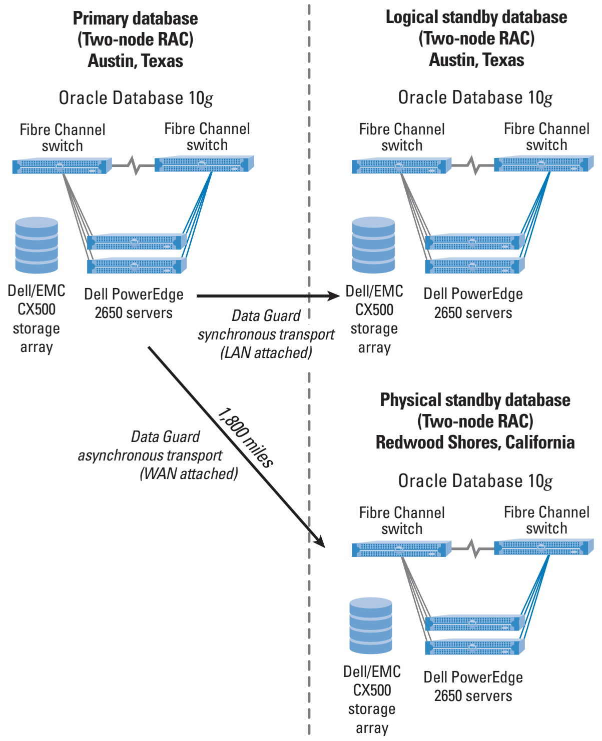

Primary database
(Two-node RAC)
Austin, Texas
Oracle Database 10g
Fibre Channel switch
Fibre Channel switch
Dell/EMC CX500 storage array
Dell PowerEdge 2650 servers
Logical standby database
(Two-node RAC)
Austin, Texas
Oracle Database 10g
Fibre Channel switch
Fibre Channel switch
Dell/EMC CX500 storage array
Dell PowerEdge 2650 servers
Data Guard synchronous transport (LAN attached)
Data Guard asynchronous transport (WAN attached)
1,800 miles
Physical standby database
(Two-node RAC)
Redwood Shores, California
Oracle Database 10g
Fibre Channel switch
Fibre Channel switch
Dell/EMC CX500 storage array
Dell PowerEdge 2650 servers

Figure 2. Dell and Oracle MAA configuration

Oracle database considerations

When building an MAA environment, administrators organize the major tasks into four main categories: preparing the cluster for the primary and standby sites, building and configuring the primary database, building and configuring the physical standby database, and building and configuring the logical standby database. The following sections provide an overview of steps that were completed in each category to implement the example Dell and Oracle configuration described in this article.[10]

Preparing the cluster to run RAC. Dell and Oracle engineers completed the following steps on each node in the server cluster for the primary database site in Austin, the physical standby

[6] For more information about MAA best-practices publications, visit www.oracle.com/technology/deploy/availability/htdocs/maa.htm.

[7] For more information about Oracle RAC, visit www.oracle.com/technology/products/database/clustering/index.html.

[8] For more information about Oracle Data Guard, visit www.oracle.com/technology/deploy/availability/htdocs/DataGuardOverview.html.

[9] For a comprehensive list of Dell and Oracle supported and validated configurations, visit www.dell.com/downloads/global/solutions/Dell-Oracle10g%20Software%20Product%20Matrix%20version%201.1.pdf.

[10] For detailed MAA deployment information, visit www.oracle.com/technology/deploy/availability/htdocs/maa.htm.

database site in Redwood Shores, and the logical standby database site in Austin:

- Installed required OS packages
- Created OS groups and users
- Configured system networking
- Configured Secure Shell (SSH) and Secure Copy (SCP)
- Set OS kernel parameters
- Created the Oracle Cluster Registry (OCR) and voting disks[11]
- Installed Cluster Ready Services (CRS)
- Installed Oracle Database 10*g* software

Creating the RAC primary database. Once the cluster was operational and all Oracle software had been installed, Dell and Oracle engineers created and configured the primary database at the Austin site as follows:

- Created a RAC database using Oracle Database Configuration Assistant (DBCA)
- Configured the database for MAA best practices, which included setting up multiple control files, setting up multiple online redo log groups, enabling archivelog mode, using automatic undo management, and using locally managed table spaces
- Set appropriate initialization parameters
- Enabled Flashback Database

Creating the RAC physical standby database. Once the primary database was configured, Dell and Oracle engineers created the physical standby database at the Redwood Shores site as follows:

- Performed an online backup of the primary database using RMAN
- Configured Oracle Net Services on each node of the standby database
- Created the standby database from the previously performed backup
- Set appropriate initialization parameters for the standby database
- Added the standby database and instances into the OCR
- Configured the primary database for redo transport
- Verified that the standby database was receiving redo generated by the primary database

Creating the RAC logical standby database. To support a logical standby database in the Data Guard configuration, Dell and Oracle engineers created the logical standby database at the Austin site as follows, taking care to ensure that the tables in the primary database included only data types that were supported by Data Guard SQL Apply:[12]

- Prepared the primary database to support a logical standby database
- Created a physical standby database by following the steps recommended in the "Creating the RAC physical standby database" section in this article
- Created a logical standby control file at the primary database
- Started up and mounted the standby database using the control file
- Activated the logical standby database and assigned it a unique name using the Oracle DBNEWID utility
- Changed the database type in the OCR to indicate it is a logical standby database
- Started SQL Apply and verified that the logical standby database was performing properly

Network configuration implications for Data Guard

For the Data Guard configuration described in this study, Dell and Oracle engineers created the logical standby database in the same data center as the primary database, and connected the logical standby database to the primary database over a LAN. The physical standby database was located 1,800 miles away, connected to the primary database over a WAN.

The objective behind such a configuration, which involved two standby databases, was to use the local logical standby database for reporting activities while enabling the remote physical standby database to be used for recovery from disasters at the primary site. Because an Ethernet LAN enables high reliability and low latency, the logical standby destination was configured using the attributes `LGWR SYNC AFFIRM` on the primary database. However, in view of the possible latencies associated with an IP-based WAN, the physical standby destination was configured with the attributes `LGWR ASYNC`. This asynchronous configuration helped minimize the impact of latency on production throughput while helping to contain the exposure to data loss in the event of a disaster.[13]

Demonstration of MAA capabilities in a high-availability implementation

After the databases and hardware were configured at the sites in Austin and Redwood Shores, Dell and Oracle engineers tested and validated the example configuration in various ways. This

[11] For more information about the OCR and voting disks, visit download-west.oracle.com/docs/cd/B14117_01/rac.101/b10765/toc.htm.

[12] For a complete list of data types not supported by Data Guard SQL Apply, visit download-west.oracle.com/docs/cd/B14117_01/server.101/b10823/toc.htm.

[13] For a comprehensive discussion of network configuration best practices for Data Guard, visit www.oracle.com/technology/deploy/availability/htdocs/maa.htm.

section highlights the major validation steps completed by the Dell and Oracle team.

Measuring load generation and network throughput

To properly demonstrate the MAA environment, Dell and Oracle engineers generated a transaction load in the primary database using a simple SQL*Loader script that inserted 20 million records into a test table. This generated redo on the primary database, and Data Guard transmitted that redo to the standby databases to keep the standby databases synchronized with the primary database. Because the redo was transmitted over the network, engineers verified that the network infrastructure was adequate to support the Data Guard configuration. This was accomplished by determining the peak database redo generation rate and comparing that rate against the network throughput. The peak redo generation rate was observed by generating Oracle Automatic Workload Repository (AWR) reports.[14] The effective network throughput rate was measured using a utility called Iperf.[15] Using these tools, engineers measured the throughput for the WAN between Austin and Redwood Shores. This approach can help an organization evaluate whether the network in its own specific Data Guard configuration can provide adequate bandwidth to support the transaction load of a particular primary database.

The objective of MAA—Oracle's blueprint for Oracle high-availability technologies—is to remove the complexity of designing an optimal high-availability architecture and to help maximize systems availability.

Using the physical standby database in read-only mode

The physical standby database can be used to perform read-only reporting in addition to providing disaster protection for the primary database. In the example implementation described in this article, engineers enabled the read-only reporting capability by temporarily suspending Redo Apply operations on the physical standby database at Redwood Shores and then opening the database in read-only mode. Once the physical standby database was open in read-only mode, engineers ran simple queries on the tables in the physical standby database to verify that the standby database was functioning properly.

After they completed this exercise, Dell and Oracle engineers restarted Redo Apply operations in the physical standby database. The Redo Apply operations automatically applied to the physical standby database all the redo that had accumulated in the physical standby database server while the physical standby database was open in read-only mode—thereby bringing the physical standby database up-to-date with the primary database.

Using the logical standby database for reporting

The logical standby database can provide advanced reporting capabilities because it is designed to be opened in read-write mode while changes from the primary database are applied to it. In the example configuration described in this article, such advanced reporting capabilities indicate that the logical standby database at the Austin site could be used as a real-time reporting system when the logical standby destination is configured with `LGWR SYNC` in the primary database and SQL Apply is running in Real-Time Apply mode[16] on the logical standby database.

Dell and Oracle engineers validated the real-time reporting capabilities of the example configuration described in this article by verifying that updates occurring on the primary database were instantaneously observed on the logical standby database. To help ensure a transactionally consistent view of the data as it was applied on the logical standby database, the SQL Apply parameter `TRANSACTION_CONSISTENCY` was set to `FULL`.

Performing a switchover

The fundamental benefit of a Data Guard configuration is that any standby database can be chosen to take over as the primary database. This approach enables organizations to resume business operations without incurring significant downtime if an outage affects the primary data center. Data Guard offers two types of role transitions: *switchover,* to be invoked for planned maintenance, and *failover,* to be invoked after unplanned outages or disasters at the primary site.

To simulate a planned outage using the example implementation and demonstrate the ease with which Data Guard can perform such role transitions, Dell and Oracle engineers executed a switchover across the WAN between the primary database in Austin and the physical standby database in Redwood Shores. Before performing the switchover, the test team verified that the standby database had recovered all available redo and that the managed recovery process was running on the physical standby database. Also, all primary and standby instances, with the exception of the ones in which the switchover was performed, were shut down.[17]

[14] For more information about AWR, visit download-west.oracle.com/docs/cd/B14117_01/server.101/b10739/toc.htm.

[15] For more information about Iperf, visit dast.nlanr.net/Projects/Iperf.

[16] For more information about Real-Time Apply, visit download-west.oracle.com/docs/cd/B14117_01/server.101/b10823/toc.htm.

[17] For more information about Data Guard switchover and failover best practices, visit www.oracle.com/technology/deploy/availability/htdocs/maa.htm.

The switchover commands—one command converts the primary database to a standby database and another command converts the standby database to the primary database—were respectively entered on the remaining instances of the primary and standby databases. The `database_role` column in the `v$database` view confirmed the status of the databases after the switchover was complete. The remaining instances of the new primary database and the new standby database were restarted, after which the database at the Redwood Shores location became the full-fledged primary database. Applications could then be directed to this new primary database in Redwood Shores, which would, in turn, start generating redo and transmit this redo to the two standby databases at the Austin site.

No data was lost in the process of the database switchover. Because the Real-Time Apply feature of Oracle Database 10*g* was used in the example configuration, the Data Guard switchover process did not have to wait for accumulated archived logs to be applied, as was the case in previous Oracle database releases. In addition, appropriate settings had been configured in all three databases before the switchover occurred, in anticipation of possible role transitions. For example, engineers set up the Oracle Database 10*g* `VALID_FOR` attributes for the `log_archive_dest_n` parameters and turned on supplemental logging at the physical standby database. Thus, the logical standby database in Austin was automatically brought along as a logical standby database to the new primary database in Redwood Shores.

Note: To revert to the original configuration following a switchover—that is, to perform a *switchback*—administrators can simply perform another switchover operation.

Monitoring Data Guard performance

The performance of a Data Guard configuration depends on whether the Redo Apply and SQL Apply operations on the physical and logical standby databases are able to keep up with the redo generation rates of the primary database.

The Oracle database offers various views that allow administrators to track the progress of the standby databases. Some relevant views to monitor the progress of the physical standby database are `v$managed_standby` and `v$dataguard_status`. Similarly, the logical standby database can be monitored using the views `v$logstdby` and `dba_logstdby_progress`.

The information provided by these views was used in the study described in this article to help engineers determine that the redo data generated by the primary database was indeed being transmitted to the standby servers across the network, and applied to the standby databases according to the configuration parameters set forth by the project team.

Understanding redo transport and networking best practices

The extent of data protection in a Data Guard configuration is controlled by the Data Guard protection mode in effect, which in turn determines whether redo is sent to the standby database synchronously or asynchronously over the network. To maintain high primary redo generation rates with the least impact on application throughput at the primary database, administrators should allow sufficient low-latency network bandwidth between the primary and standby sites and properly configure relevant OS network settings (such as TCP send and receive buffer sizes). In general, MAA best practices recommend `LGWR SYNC` (Maximum Protection or Maximum Availability protection mode) for low-latency LAN or metropolitan area network (MAN) environments in which the standby site may be located up to a few hundred miles from the primary site. `LGWR ASYNC` or ARCH-based redo transport (Maximum Performance protection mode) is recommended for more distant standby databases connected by an IP-based WAN.

Characteristics of a highly resilient IT architecture

A global enterprise must be based on a highly resilient IT architecture to provide mission-critical levels of service to customers and stakeholders around the world. Such an architecture must be complete, integrated, easy to manage, and flexible enough to serve multiple purposes. At the same time, the technology to implement a highly resilient IT architecture should be cost-effective to enable businesses to derive optimal value from their IT investments.

Oracle Database 10*g* offers an integrated suite of high-availability capabilities that can help meet these demanding business and technical requirements. The implementation discussed in this article demonstrates the ease with which Oracle Database 10*g* MAA can be deployed over a WAN on Dell servers and a Dell/EMC storage platform to provide a robust degree of high availability. Using the configuration best practices discussed in this article, enterprises can deploy a combined Dell and Oracle platform to build an end-to-end, highly available and scalable infrastructure.

Michael Smith is a senior member of technical staff with the Oracle Database High Availability Group. His principal focus is Oracle Data Guard. He has a bachelor's degree in Computer Science.

Paul Rad is a senior software engineer in the Dell Database and Application Engineering Department of the Dell Product Group. He has master's degrees in both Computer Engineering and Computer Science from The University of Texas at San Antonio.

Ashish Ray is a group product manager with the Oracle Database High Availability Group. His principal focus is Oracle Data Guard. He has bachelor's and master's degrees in Computer Science, and an M.B.A.

Using Intel Multi-Port Server Adapters to Enable Virtual Infrastructure in the Data Center

A virtual computing infrastructure can provide IT staff with the capability to improve hardware resource utilization, thereby helping to reduce costs and streamline IT administration. A virtual computing environment also helps keep mission-critical applications available and allows administrators to respond quickly to ever-changing business needs. To take best advantage of the benefits of virtualization, IT organizations also require critical hardware components that are designed to enable high availability, reliability, and performance: multi-port server adapters.

BY BILL HENDERSON AND TRACY D. EDWARDS

Related Categories:

Data networking

Dell PowerEdge servers

Intel networking

Network fabric

Network interface card (NIC)

Server consolidation

Virtualization

VMware

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

IT organizations today face unprecedented requirements to keep pace with rapidly evolving business needs while maintaining continuous availability of mission-critical applications and keeping costs low. One consequence of the scramble to provide fast and flexible service is the tendency toward a sprawl of underutilized servers in the data center and an IT staff that is continually reacting to changing business conditions instead of proactively planning for growth. To rein in IT costs and keep enterprise applications running reliably without service interruptions, organizations are exploring ways to make their existing hardware resources do more.

Hardware virtualization offers the basis for a flexible, low-cost IT infrastructure that can provide the capability to respond immediately to changing business needs. By enabling a reduction in the total number of physical servers in the IT environment, a virtual computing infrastructure can help streamline systems management and reduce total cost of ownership (TCO). At the same time, a virtual infrastructure can allow administrators to move application workloads easily from one physical server to another, facilitating seamless business continuity and disaster recovery.

Key to a virtual computing infrastructure is a software platform that provides a virtualization layer designed to decouple application workloads from the underlying hardware of the physical server. In this way, virtualization software can provide a set of virtual computing, memory, networking, and storage resources to each application. Virtualization software is designed to enable a single physical server to be divided into several independent virtual machines (VMs), each of which can host a separate “guest” OS and associated applications in complete isolation from other VMs on the server (see Figure 1). The virtual infrastructure approach enables administrators to manage and optimize resources transparently across the data center.

The capability to run multiple VMs simultaneously on the same physical server enables enterprises to consolidate workloads from several separate physical servers onto one server, helping to reduce the number of physical servers required for a given workload. Unlike physical servers, VMs can be deployed in a matter of minutes and moved from one physical server to another without reconfiguring the OS, hardware, or applications. As a result, administrators can provision services quickly and allocate

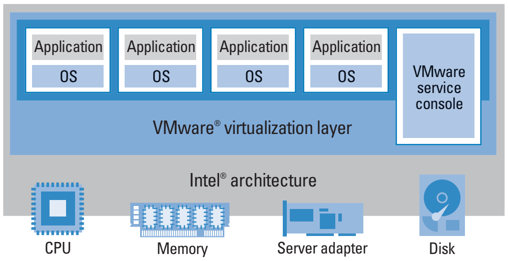


Figure 1. Virtual infrastructure approach—providing a layer of abstraction between VMs and underlying physical hardware

resources cost-effectively to business units as necessary. In this way, virtualization can help IT administrators prevent unplanned downtime and increase overall application availability.

Building a virtual infrastructure with VMware software

VMware software is designed to create a virtual computing infrastructure that allows administrators to deliver IT services quickly and flexibly. At the same time, VMware virtualization software can help reduce the cost of computing by allowing administrators to deploy additional VMs on the same physical server—thereby improving server utilization and simplifying systems management. The building blocks of a VMware-based virtual computing infrastructure include VMware® GSX Server™ and VMware ESX Server™ virtualization software, VMware VirtualCenter management software, and VMware VMotion™ technology.

VMware GSX Server and ESX Server. VMware GSX Server and ESX Server run on Intel® Xeon™ processor–based platforms, providing a virtualization layer that allows multiple VMs to coexist in isolation from one another on a single physical server. GSX Server and ESX Server software typically enable organizations to run from one to eight VMs per processor on two-processor and four-processor systems such as Dell™ PowerEdge™ 1850, PowerEdge 2850, PowerEdge 6650, and PowerEdge 6850 servers as well as the PowerEdge 1855 blade server. In addition, ESX Server works with VirtualCenter and VMotion to provide additional IT management capabilities (see the “VMware VirtualCenter and VMotion technology” section in this article).

To take best advantage of the capabilities of VMware virtualization software when consolidating workloads from several physical servers onto one physical server, administrators must carefully manage the capacity of network adapters on physical servers that host multiple VMs. High-speed Intel Gigabit[1] Ethernet server adapters can play an essential role in enabling the virtual computing infrastructure because multiple VMs running on one physical server create a need for additional server ports. Intel multi-port server adapters address this need by providing additional ports to support virtualization in slot-constrained servers. Multi-port server adapters also give IT administrators the flexibility to configure redundant ports to help improve throughput and reliability.

VMware VirtualCenter and VMotion technology. Designed to give administrators full control over ESX Server–based resources in the data center, VMware VirtualCenter management software works together with VMware VMotion technology, which supports dynamic migration of applications running on VMs. VirtualCenter can enhance the management of virtualized Intel architecture–based environments whether they use Microsoft® Windows®, Novell® NetWare®, or Linux® operating systems. By treating the Intel processor–based hardware on physical servers as a single logical pool of computing resources, VirtualCenter helps optimize resource allocation, enables centralized monitoring of systems for availability and performance, and deploys VMs using standardized templates to ensure consistent server images.

VMware VMotion technology allows administrators to use VirtualCenter to quickly and easily migrate an active VM from one physical server to another without service interruption, making dynamic workload balancing and zero-downtime hardware maintenance possible (see Figure 2). As a result, VMotion enables fast reconfiguration and optimization of resources across the virtual infrastructure while maintaining continuous application availability.

Enhancing data center efficiency and flexibility

VMware virtual infrastructure software can benefit enterprises by enabling a responsive, cost-effective IT infrastructure, as follows:

- **Low TCO through server consolidation.** VMware software can help provide significant savings in capital hardware expenditures and ongoing operating costs by enabling server consolidation. For example, using VMware ESX Server, administrators can host 16 separate servers (operating systems and associated applications) on a two-processor system and more than 32 separate servers on a four-processor system. This approach enables organizations to minimize hardware requirements by consolidating applications and services that currently run under various operating systems onto fewer, highly scalable and reliable enterprise-class servers, including blade servers.

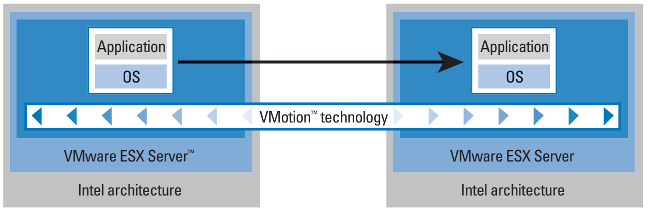


Figure 2. VirtualCenter migration of a running VM from one physical server to another

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

- **Streamlined systems management.** VMware-enabled server consolidation helps enhance data center operations by minimizing the administrative effort required to deploy, manage, and update servers in the virtual computing environment. IT managers can put unused server capacity to work simply by balancing workloads across physical servers and adjusting the resources dedicated to each VM.
- **Quick response to business needs.** VMware VirtualCenter management capabilities are designed to provision servers in minutes, allowing IT administrators to quickly respond to a variety of demands, including increased demand for capacity, new requests for IT services, and the need for hardware and performance upgrades. In addition, VMware virtualization software can facilitate fast, cost-effective time to market by allowing administrators to set up large-scale development and test environments on fewer physical servers running VMs as compared to physical servers not running VMs—enabling comprehensive testing without incurring the cost, time, and resource burden normally required to provision and configure a large number of physical servers.
- **Cost-effective business continuity and disaster recovery.** Because VMotion technology allows live applications running in a VM to be moved from one physical server to another, administrators can perform maintenance on mission-critical servers without a disruption in service—thus dramatically improving application availability. In addition, VMs can be clustered to further increase availability by providing automatic failover from one server to another. VMware software can also help administrators implement cost-effective disaster recovery solutions. For example, if a disaster occurs, administrators can recover production workloads running on VMs in minutes simply by copying the VM images to a physical server at the disaster recovery site and then restarting the VMs on that server.

Providing optimal network port capacity in a virtual computing environment

To achieve the preceding IT benefits using a virtualization platform such as VMware software, IT organizations must equip physical servers running VMs with the appropriate number and type of network connections. Although frequently taken for granted, adequate network port capacity is essential to enabling optimal functionality in a virtual computing environment.

For example, best practices for VMware ESX Server recommend a minimum of three network adapters—one for the VMware service console, to enable system administration; one for VMotion, to enable dynamic workload balancing; and at least one for the VMs and their applications. However, providing this number of adapters can be a challenge if physical servers are constrained by an insufficient number of Peripheral Component Interconnect (PCI) slots.

In addition, administrators must mitigate reliability risks. On a server running multiple VMs, a port failure can be time-consuming to fix and can entail a costly business interruption—unless redundant network connections are used to enhance the reliability of servers hosting mission-critical applications. For enhanced availability in a virtual computing infrastructure, up to six ports may be required per physical server.

IT administrators also need the capability to segment traffic to avoid bottlenecks in the network. The flexibility to assign network ports to servers and partitions enables administrators to quickly increase performance during peak times. As server workloads grow, enterprises need network port adapters that provide the throughput required to accommodate increased requests for data. Such requirements consume valuable server ports.

Using multi-port Intel PRO Server Adapters to enhance reliability and throughput

Multi-port server adapters such as Intel® PRO/1000 Dual Port and Quad Port Server Adapters are designed to fit additional ports into a small form factor, conserving valuable PCI slots in servers while helping to eliminate network bottlenecks in connection-dense virtual computing environments. Intel multi-port server adapters can also help organizations migrate easily and cost-effectively to Gigabit Ethernet because they are designed with an integrated Gigabit Ethernet controller chip that enables high performance and reliability as well as low power consumption. Optimized for Intel Xeon processor–based servers, these adapters help provide the foundation for a flexible and reliable virtual computing infrastructure.

Intel multi-port server adapters provide two or four Gigabit Ethernet connections in a single PCI card. ESX Server–based configurations running VirtualCenter require one dedicated port for the VMware service console and another dedicated port for VMotion. In addition, best practices for ESX Server–based configurations recommend that administrators dedicate one or more server adapters to each VM (see Figure 3). Intel multi-port server adapters can provide the ports to meet this recommendation.

Advanced server features. In addition to providing needed connectivity, Intel multi-port Gigabit Ethernet adapters can help prevent network downtime and enable maximum processor utilization. For example, Intel multi-port adapters include support for advanced server features including adapter fault tolerance, which is designed to provide redundant network links for server failover. In addition, Intel multi-port server adapters support adaptive load balancing and link aggregation, which are designed to enhance scalability and throughput. PCI Hot Plug and ActivePCI features help keep systems up and running, while interrupt moderation can significantly enhance processor utilization.

Extensive PCI slot compatibility. With a flexible design that fits almost any type of PCI bus, Intel PRO Dual Port and Quad Port

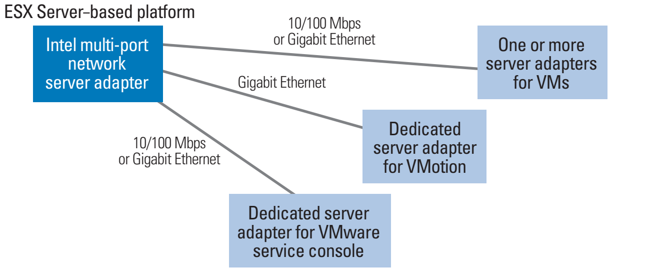

ESX Server–based platform
Intel multi-port network server adapter
10/100 Mbps or Gigabit Ethernet
One or more server adapters for VMs
Gigabit Ethernet
Dedicated server adapter for VMotion
10/100 Mbps or Gigabit Ethernet
Dedicated server adapter for VMware service console

Figure 3. Recommended network connections on a server running ESX Server

Server Adapters include standards-based management features and wide network OS support to help ensure extensive compatibility with the latest server and networking environments. Intel adapters are compatible with full-height and low-profile PCI slots. As a result, administrators can easily replace the standard, full-height bracket with the shorter, low-profile bracket for installation in high-density servers that have low-profile PCI slots.

High reliability through redundancy. Intel® Advanced Network Services (ANS) software allows administrators to maximize uptime with redundant, teamed connections. Switch fault tolerance and test switch configuration features help administrators verify compatibility and enhance uptime. Intel multi-port adapters can be configured to automatically switch to a secondary link when a server's primary link fails. Server performance can be further enhanced by teaming connections on adapters with each other, with other Intel PRO adapters or Intel LAN on Motherboard (LOM) adapters, or with third-party adapters using Intel ANS features. The teaming approach can help enterprises scale up to Gigabit Ethernet and provide redundancy for server failover capability. Advanced cable diagnostics can dynamically test and report network problems such as interrupts and automatically compensate for cable problems.

Server expandability and business continuity. Proactively installing an Intel multi-port server adapter with extra ports can support server expandability by helping enterprises avoid the cost of taking a server offline to install an additional network port when adding servers or VMs. Intel multi-port adapters also help enable business continuity by allowing IT administrators to dedicate ports for remote storage and management including storage area network (SAN), network attached storage (NAS), and direct attached storage (DAS).

Network segmentation. To support unpredictable network demands such as heavy traffic on Web servers and intranets, many IT organizations are segmenting enterprise networks. Segmentation is designed to help enhance data security and uptime for each network. For example, within dedicated subsystems, hardware and software operating environments cannot be modified without proper authentication. Employing multi-port network adapters can provide the platform flexibility to respond to evolving needs for segmentation and enable the connection to multiple switches to segment traffic on a network.

Cabling advantages. Intel offers dual-port and quad-port adapters for both copper and fiber-optic networks. Both types of adapters use a common driver technology—Intel® SingleDriver™ technology—for Gigabit Ethernet, which helps reduce IT complexity. Copper Intel Gigabit Ethernet adapters support cost-effective 10/100/1000 Mbps transmission rates over existing Category 5 cabling. IT administrators can upgrade Fast Ethernet connections to Gigabit Ethernet using Intel PRO adapters. This added flexibility can help reduce training costs associated with the upgrade and expedite the Gigabit Ethernet rollout.

Centralized remote management. Intel PRO/1000 Dual Port and Quad Port Server Adapters are designed to support standard management protocols. These protocols include Wired for Management (WfM), Microsoft Remote Installation Service (RIS), Simple Network Management Protocol (SNMP), and Desktop Management Interface (DMI).

Making the most of IT resources through virtualization

To maximize return on investment and help keep TCO low, enterprises must capitalize on IT resources already in place. A virtual computing infrastructure can be a powerful enabler for consolidating servers, running applications in multiple-OS environments, simplifying administration, and minimizing data center operating costs. VMware software running on Intel Xeon processor–based servers is designed to enable a robust virtualization platform, particularly when combined with Intel multi-port server adapters that provide the connections needed to optimize the virtual IT infrastructure. The approach discussed in this article can help IT organizations reach their common goal: serving business needs in the most efficient and responsive way possible.

Bill Henderson is a senior systems engineer for the Strategic Alliances organization at VMware, Inc. He has more than 20 years of experience in various roles in the computer industry.

Tracy D. Edwards is a technical marketing engineer in the LAN Access Division at Intel Corporation. With more than 11 years of networking experience, Tracy has designed, administered, and deployed international wide area networks.

FOR MORE INFORMATION

Virtualization solutions from Intel and VMware:
www.intel.com/network/connectivity/solutions/virtualization.htm

Gigabit solutions from Intel and Dell:
www.intel.com/go/dellgig10

VMware virtual infrastructure software:
www.vmware.com

Dell and VMware:
www.vmware.com/dell
www.dell.com/vmware

Enabling Dell OpenManage Applications for Microsoft Active Directory User Authentication

Many enterprises have implemented security policies that require user authentication for directory services such as Microsoft® Active Directory® directory service. Certain components of the Dell™ OpenManage™ systems management suite—Dell OpenManage Server Administrator, Dell OpenManage IT Assistant, and Dell Remote Access Controller 4—have been enabled for Microsoft Active Directory user authentication. This article describes Dell's approach to the security integration of these components and provides implementation examples.

BY MARCOS PALACIOS, PH.D.

Related Categories:

Authentication

Dell OpenManage

Dell PowerEdge servers

Directory services

Microsoft Active Directory

Microsoft Windows

Systems management

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

A directory is a repository used to store information about a set of relevant objects. For example, a telephone directory stores information about telephone service subscribers in a given locale. In a server's file system, the directory stores information about files on the server and their respective locations and attributes. In a typical distributed enterprise computing system, many types of relevant objects need to be stored in a directory—including user account names, application servers, Web servers, printers, fax servers, and other objects. End users want to easily and quickly find the objects they need, while IT administrators are concerned with managing and securing access to the various objects to comply with their organization's security and usage policies.

Directory service capabilities and features

A robust directory service is one of the most important components of an extended computer system. Users and administrators frequently do not know the exact names of the objects they need to access. They may, however, know one or more attributes of these objects.

A directory service allows a user to find any object based on one or more of its attributes by querying the directory to obtain a list of objects that match the attributes. For example, an administrator might pose the query: "Find all duplex printers in building 26." In addition, a directory service can perform the following actions:

- Enforce security defined by administrators to help keep information safe from intruders
- Distribute a directory across several computers in a network
- Replicate a directory to make it available to more users and to keep it highly available
- Partition a directory into multiple stores to allow the storage of very large numbers of objects

Microsoft Active Directory is the directory service included with Microsoft Windows® 2000 Server and Windows Server™ 2003 operating systems. It extends the functionality of previous Windows-based directory services and introduces additional features. Active Directory is secure, distributed, partitioned, replicated, and highly scalable. It is designed to work well in any size installation, from a single server hosting a few hundred objects to thousands of servers hosting millions of objects. Active Directory's feature set helps ease navigation and management of large amounts of information—helping to save time and improve productivity for both administrators and end users.

Active Directory is a distributed database. The rules for the database are defined by the database schema, which is a collection of attributes and classes. An example class is the User class. Examples of attributes from the User class are First Name, Last Name, and Phone Number.

Dell OpenManage support for Microsoft Active Directory

Several products in the Dell OpenManage suite now support Active Directory user authentication: Dell OpenManage Server Administrator (OMSA), Dell OpenManage IT Assistant (ITA), and the Dell Remote Access Controller 4 (DRAC 4). A major advantage of Dell's support for Active Directory is that supported Dell products use the same authorization and authentication schemas and associated security configuration user interfaces as Active Directory, which enhances the end user's experience by helping to reduce the complexity associated with directory services security. Figure 1 shows how the applications comprising the Dell OpenManage suite can integrate with Microsoft Active Directory directory services security.

Dell schema extensions for Active Directory enablement

Dell has extended the Active Directory schema by adding attributes and classes to represent Dell OpenManage objects. By doing so, Dell has tailored Active Directory to meet Dell OpenManage user authentication and authorization needs.

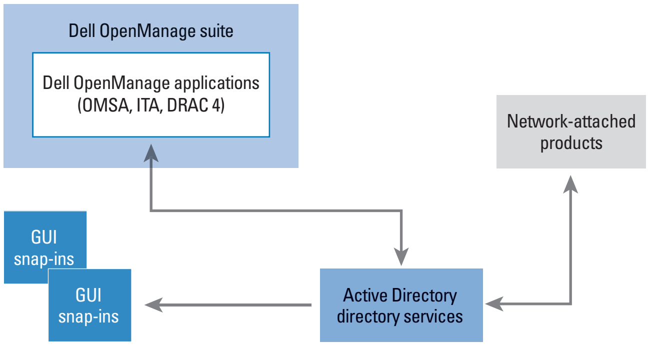

Dell OpenManage suite
Dell OpenManage applications (OMSA, ITA, DRAC 4)
Network-attached products
GUI snap-ins
GUI snap-ins
Active Directory directory services

Figure 1. Dell OpenManage integration with Microsoft Active Directory security

Dell has defined a group of Active Directory objects that can be configured depending on an organization's IT environment: an Association object, a Device object, an Application object, and a Privilege object.

- **Association object:** Links together users or groups possessing a specific set of privileges to one or more Dell OpenManage devices or applications
- **Device object:** Represents a Dell remote access controller (RAC) device
- **Application object:** Represents either a Dell OMSA application or a Dell ITA application
- **Privilege object:** Lists which privileges are granted to users depending on the Dell OpenManage application or device

One Application object or Device object must exist in the Active Directory database for each Dell OpenManage application or device to be managed. Administrators can create as many Privilege objects as there are levels of users in the organization. Similarly, administrators can create as many Association objects as there are relationships between users and the applications or devices to be implemented in the environment.

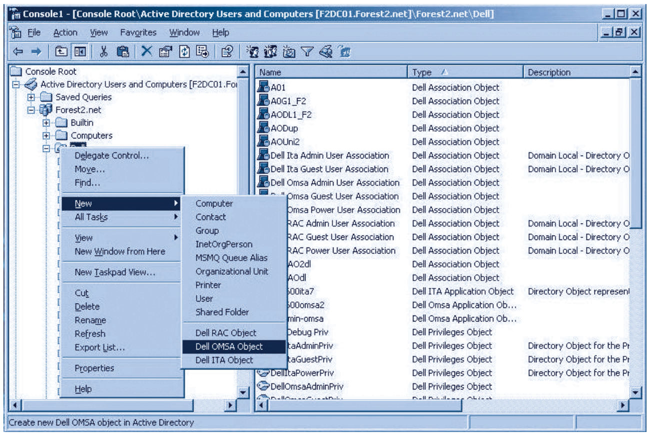
To help administrators modify the Active Directory schema, Dell provides both a wizard installation utility—called the Schema Extender Utility—and a command-line Lightweight Directory Access Protocol (LDAP) Data Interchange Format (LDIF) configuration file for each Dell OpenManage application and device to be Active Directory enabled. The Schema Extender Utility walks the administrator through the process of extending the schema. When executed, the utility is designed to display the results of each of the attributes and classes added to the schema and a message acknowledging that the Active Directory objects have been successfully added. If another administrator has already run the Schema Extender Utility and made the changes to the schema, then the utility generates a message indicating that the Active Directory objects already exist.

A robust directory is one of the most important components of an extended computer system.

The LDIF configuration file is for advanced Active Directory administrators who want to see the specific modifications that they have configured before these changes are made to the Active Directory schema. Microsoft Windows operating systems provide a utility called ldifde.exe that is used to run the LDIF configuration file.

Extension to the Microsoft Management Console snap-in

After extending the schema, administrators must extend the Microsoft Management Console (MMC) Active Directory Users and


Console1 - [Console Root\Active Directory Users and Computers [F2DC01.Forest2.net]\Forest2.net\Dell]

Name	Type	Description
A01	Dell Association Object	
A0G1_F2	Dell Association Object	
AODL1_F2	Dell Association Object	
AODup	Dell Association Object	
AOUni2	Dell Association Object	
Dell Ita Admin User Association	Dell Association Object	Domain Local - Directory O
Dell Ita Guest User Association	Dell Association Object	Domain Local - Directory O
Dell Omsa Admin User Association	Dell Association Object	
Omsa Guest User Association	Dell Association Object	
Omsa Power User Association	Dell Association Object	
RAC Admin User Association	Dell Association Object	Domain Local - Directory O
RAC Guest User Association	Dell Association Object	Domain Local - Directory O
RAC Power User Association	Dell Association Object	Domain Local - Directory O
AO2dl	Dell Association Object	
AOdl	Dell Association Object	
600ita7	Dell ITA Application Object	Directory Object represen
600omsa2	Dell Omsa Application Ob...	
min-omsa	Dell Omsa Application Ob...	
Debug Priv	Dell Privileges Object	
taAdminPriv	Dell Privileges Object	Directory Object for the Pr
taGuestPriv	Dell Privileges Object	Directory Object for the Pr
DellItaPowerPriv	Dell Privileges Object	Directory Object for the Pr
DellOmsaAdminPriv	Dell Privileges Object	

Create new Dell OMSA object in Active Directory

Figure 2. Adding an Application object to the Microsoft Active Directory database

Computers snap-in. This allows administrators to create the Active Directory objects needed to manage RAC devices, OMSA applications, and ITA applications. Dell provides a wizard installation application that walks administrators through modifying the MMC snap-in. Administrators must install this utility on each server that is used to access the MMC Active Directory Users and Computers snap-in utility.

Every Active Directory domain contains a set of containers that are created during the installation of Active Directory. Organizational units are Active Directory containers into which administrators can place users, groups, computers, and other organizational units. From the menu associated with a container or organizational unit in the Active Directory Users and Computers console, the snap-in extensions allow administrators to create Active Directory objects by selecting either New > Dell RAC Object; New > Dell OMSA Object; or New > Dell ITA Object (see Figure 2). Administrators can add these objects by right-clicking on the container or organizational unit to which they want to add the object. The menu structure shown in Figure 2 appears if the container or organizational unit supports Dell objects.

Dell OpenManage and Active Directory integration

Once the extensions are installed in the Active Directory database, new Active Directory objects for Dell OpenManage integration can be added. Administrators can add a Device object, for example, using the following steps to create each device and application:

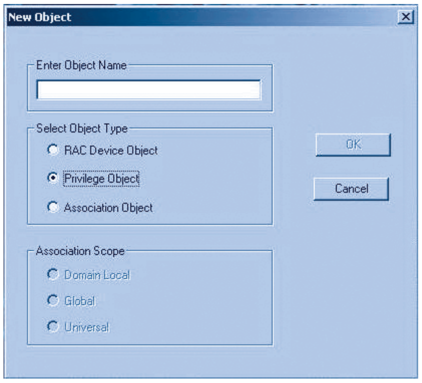
1. Select New > Dell RAC Object from the Active Directory console menu, and a dialog box similar to the one shown in Figure 3 is displayed.
2. Choose the object name, type of object, and scope characteristics of the object. The object will then be created and added to the container from which the command was initiated.

To create an Association object, administrators must choose the association scope that applies to the type of object added. The association scope is the security group type that the association object will have. The Association object is derived from a security group and must contain a group type. Universal Association objects can be created only when the Active Directory domain is functioning in native mode or higher.

Dell has tailored Active Directory to meet Dell OpenManage user authentication and authorization needs.

By right-clicking on an Association object's Properties page, administrators can add the desired users or groups, a Privilege object, and Dell products to the association. Figure 4 shows a user called Administrator being added to an Association object.

Similarly, by clicking on the Privilege Object tab, administrators can add the Privilege object to the association that defines the user's or group's privileges when authenticating to a RAC device or Dell OpenManage application. *Note:* Only one Privilege object can be added to an Association object. Additionally, by clicking on the Products tab, administrators can add one or more Dell products to the association. These products specify the RAC devices or Dell OpenManage applications that are available for the defined users or groups. Administrators can add multiple Dell products to an Association object by using the Add button.

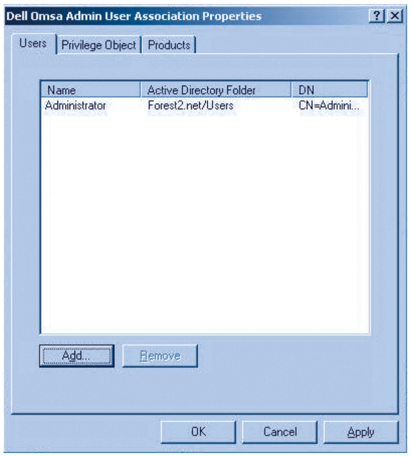

New Object

Enter Object Name

Select Object Type
- ○ RAC Device Object
- ● Privilege Object
- ○ Association Object

Association Scope
- ○ Domain Local
- ○ Global
- ○ Universal

Figure 3. Adding a Device object to the Active Directory database


Dell Omsa Admin User Association Properties

Name	Active Directory Folder	DN
Administrator	Forest2.net/Users	CN=Admini...

Figure 4. Adding a user called Administrator to the Association object

Active Directory configuration parameters for Dell products

The final step in establishing communication with the Active Directory servers is to configure the individual RAC devices and Dell OpenManage applications with specific parameters. Administrators must configure the following Active Directory settings across devices and applications using their respective interfaces:

- **Active Directory Dell product name:** This is a unique string representing the name of the Dell product. The same name should be used both for creating the Dell product in the Active Directory environment and adding the Dell product to the Dell Association object.
- **Active Directory domain name:** This is the Active Directory root domain name—for example, mydomain.com.
- **Active Directory Certificate Authority (CA) certificate:** This certificate is created from the organization's Active Directory certification authority. The certificate is downloaded to a file and then uploaded to the server where the Dell product is located.
- **Dell product certificate:** This certificate allows the Dell product to communicate securely with the Domain Name System (DNS) server to authenticate a user in the Active Directory database. The Dell product certificate is downloaded to a file and then uploaded to the Active Directory domain being accessed.

IT management advantages of Active Directory support

By delivering support for Microsoft Active Directory authentication in Dell OpenManage applications and RAC devices, Dell has enabled IT administrators to seamlessly integrate management of Dell products into the directory service that they are already using to manage other objects in their enterprise. The benefits of this standards-based approach include maximum flexibility for IT administrators in controlling access to Dell OpenManage applications, granular assignment of privileges based on the type of Dell OpenManage application being accessed, and consolidation of security processes in a central Active Directory repository rather than having to distribute these security processes among local devices.

Marcos Palacios, Ph.D., is a software quality engineer on the Dell OpenManage development team. Prior to joining Dell, he worked for BMC Software, where he specialized in software test processes and methodologies. Marcos has a Ph.D. from Texas Tech University.

FOR MORE INFORMATION

Dell OpenManage systems management:
dell.com/openmanage

Maximizing Remote Management Security on Eighth-Generation Dell PowerEdge Servers

IT administrators can take advantage of a powerful option for managing remote servers through out-of-band connections by using the Intelligent Platform Management Interface (IPMI) together with the integrated baseboard management controller in eighth-generation Dell™ PowerEdge™ servers. This article discusses key security features that are part of the IPMI 1.5 standard, and examines how the latest Dell remote access controllers can help administrators enhance remote server management.

BY CHANDRA S. MUGUNDA, WEIMIN PAN, AND HAIHONG ZHUO

Related Categories:

Baseboard management controller (BMC)

Dell OpenManage

Dell PowerEdge servers

Dell Remote Access Controller (DRAC)

Intelligent Platform Management Interface (IPMI)

Out-of-band (OOB) management

Security

Systems management

Virtual media

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

The Intelligent Platform Management Interface (IPMI) 1.5 specification describes a standard way of managing remote servers through out-of-band connections. However, while out-of-band connections provide IT administrators with a rich set of capabilities, they also introduce security challenges. To enable secure management of IPMI 1.5–compliant, eighth-generation Dell PowerEdge servers using out-of-band connections, administrators must properly configure the integrated baseboard management controller (BMC). This article explains how administrators can configure security features of the Dell Remote Access Controller 4 (DRAC 4) to help ensure tight security.

IPMI security features

When a BMC remote management connection is configured on a server—via serial, LAN, or serial over LAN (SOL) links—an application or utility that complies with the IPMI 1.5 specification can access the server through that connection. Although the IPMI specification allows the Anonymous setting to be enabled by default, eighth-generation Dell PowerEdge servers are shipped with the Anonymous setting disabled to help protect against potential security breaches.[1]

IPMI 1.5 helps provide security through user authentication; the BMC maintains a local database of remote access users and their privileges. Individual users in the BMC local user database are assigned a privilege limit that dictates the type of rights they have on the BMC. Administrators can use the Dell OpenManage™ Server Administrator (OMSA) interface to manage BMC user accounts (see Figure 1).

Access to servers can be restricted through connection-level, or *channel-level*, privileges; through user-level

[1] For more information about how to configure Dell PowerEdge server BMCs to enable server management through supported out-of-band connections, see "Remote Management with the Baseboard Management Controller in Eighth-Generation Dell PowerEdge Servers" by Haihong Zhuo; Jianwen Yin, Ph.D.; and Anil V. Rao; in *Dell Power Solutions,* October 2004.

privileges; or both. Each channel can be limited to operate at one of three different privilege levels: User, Operator, or Administrator. Similarly, each user can be created with one of these three privileges. For example, when a particular channel is limited to Operator level, only Operator-level operations can be performed on that channel.[2]

IPMI BMC authentication mechanism

Authenticated IPMI communication to the BMC is accomplished by establishing a session. Each session connection includes a user authentication phase that precedes IPMI messaging. The BMC verifies the packets that it receives. Authenticated packets are silently discarded if the authentication signature is invalid or the authentication type does not match the authentication type that was negotiated when the session was activated.

DRAC 4 security features

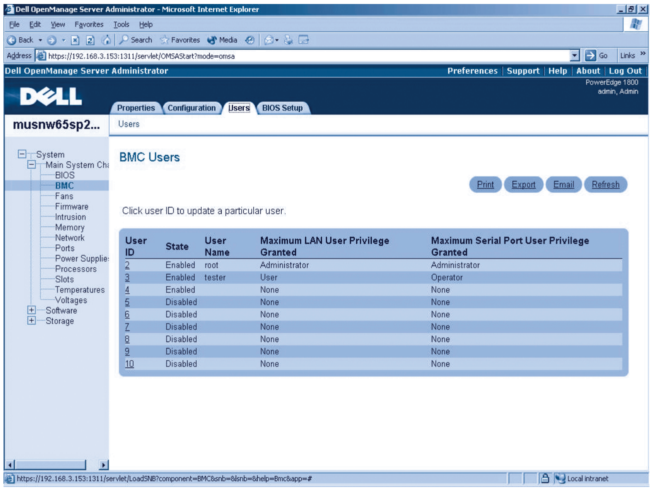
The DRAC 4 enhances current Dell remote access controller (RAC) offerings by providing features such as role-based user authentication, Racadm utility security, virtual media security, and console redirection security.

Role-based user authentication

The DRAC 4 supports privileged user-based and role-based access to a RAC device. Each DRAC 4 user entered in the RAC local user database or Microsoft® Active Directory® directory service user database is assigned a set of privileges. These privileges determine which rights the user has on the RAC device.

The DRAC 4 card supports nine privileges, which enable users to do the following:

- **RAC Login User:** Log in to the DRAC 4. Administrators can easily disable a user by removing this privilege. Removing the login privilege to disable a user is more straightforward than deleting a user. After a user's RAC Login User privilege is removed, that user still exists in the RAC or Active Directory user database. To re-enable the user at a later time, an administrator can simply grant the RAC Login User privilege again; there is no need to completely reconfigure the user in the database, as would be the case if the user were deleted.
- **RAC Card Configuration:** Change the DRAC 4 configuration—including out-of-band network interface card (NIC) configuration, Simple Network Management Protocol (SNMP) trap configuration, and Secure Sockets Layer (SSL) certificate configuration—with the exception of user configurations.


Dell OpenManage Server Administrator - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address https://192.168.3.153:1311/servlet/OMSAStart?mode=omsa
Go Links
Dell OpenManage Server Administrator
Preferences | Support | Help | About | Log Out
DELL
PowerEdge 1800
admin, Admin
Properties Configuration Users BIOS Setup
musnw65sp2...
Users
System
Main System Cha
BIOS
BMC
Fans
Firmware
Intrusion
Memory
Network
Ports
Power Supplie
Processors
Slots
Temperatures
Voltages
Software
Storage
Print Export Email Refresh
https://192.168.3.153:1311/servlet/LoadSNB?component=BMC&snb=&lsnb=&help=Bmc&app=#
Local intranet

BMC Users

Click user ID to update a particular user.

User ID	State	User Name	Maximum LAN User Privilege Granted	Maximum Serial Port User Privilege Granted
2	Enabled	root	Administrator	Administrator
3	Enabled	tester	User	Operator
4	Enabled		None	None
5	Disabled		None	None
6	Disabled		None	None
7	Disabled		None	None
8	Disabled		None	None
9	Disabled		None	None
10	Disabled		None	None

Figure 1. BMC user account management

- **RAC User Configuration:** Add or delete users, or change existing users' privileges.
- **RAC Log Clear:** Clear the system event log (SEL), RAC log, or last-crash screen log.
- **RAC Server Reset and Power-On/Off:** Perform power management operations on a system, such as reset, power up, or power down.
- **RAC Console Redirect:** Use the console redirection feature.
- **RAC Virtual Media:** Use the virtual media feature.
- **RAC Test SNMP Alert:** Set up the DRAC 4 to send a test SNMP trap alert to a preconfigured destination.
- **RAC Debug Command:** Issue debug commands. Most debug commands are used to help debug or diagnose the DRAC 4 and are normally used only by administrators or support technicians.

The DRAC 4 card supports five predefined user groups, which enable the following privileges:

- **Administrator:** A user in the Administrator group has all nine DRAC 4 privileges and can fully use all the features in the DRAC 4.
- **Power User:** A user in the Power User group has all DRAC 4 privileges except for RAC User Configuration and RAC Card Configuration. Thus, a member of the Power User group can use the DRAC 4 remote management features but cannot change any RAC configurations or user configurations.
- **Guest User:** A user in the Guest User group has only the RAC Login User privilege. A guest user can log in and view the various logs, system information, and session information.

[2] For more information about user privileges and which operations can be performed at each privilege level, refer to Appendix G of the IPMI 1.5 specification at www.intel.com/design/servers/ipmi/index.htm.

- **E-mail Alerts Only:** A user in the E-mail Alerts Only group can receive e-mail alerts but cannot log in to the DRAC 4.
- **Custom:** A user in the Custom group can be assigned any combination of privileges.

Configuring a RAC local user. Administrators can manage a DRAC 4 user through the supplied RAC GUI or the Racadm command-line interface (CLI). The Racadm utility is available on the Dell OpenManage CD. A user with the RAC User Configuration privilege—for example, a member of the Administrator group—can configure users. After logging in through a browser, the administrator must click the Configuration tab, then click the Users subtab to open a Remote Access Controller Users page. The DRAC 4 allows a maximum of 16 RAC local users. The DRAC 4 ships from Dell with a default user called root preconfigured in its first user slot; the other 15 slots are available. The root user has Administrator group privileges.

To add a user, the administrator can click on one of the available user slots to open the Add/Configure RAC User page. Three types of information need to be configured on that page: general information (including username and password), privileges, and e-mail alerts.

An administrator can place a user in a predefined group by selecting a group from the User Group list. Alternatively, the administrator can assign any set of privileges to a user by clicking the Privilege check box and placing the user in the Custom group.

If the user needs to receive e-mail alerts, the administrator must check the Enable E-mail Alert check box and configure a valid user e-mail address. The e-mail alert filter can also be configured by checking or unchecking boxes pertaining to different sensor types and severity levels. Only the alerts that pass this filter checking are sent to the user.

Users can also be added from the CLI using the Racadm utility, as follows:

```
racadm config -g cfgUserAdmin -o cfgUserAdminUserName
    -i 1 username
racadm config -g cfgUserAdmin -o cfgUserAdminPassword
    -i 1 password
racadm config -g cfgUserAdmin -o cfgUserAdminPrivilege
    -i 1 privileges
```

A user's e-mail alert can be configured from the CLI as follows:

```
racadm config -g cfgUserAdmin -o cfgUserAdminEmailEnable
    -i 1 1
racadm config -g cfgUserAdmin -o cfgUserAdminEmailAddress
    -i 1 email_address
```

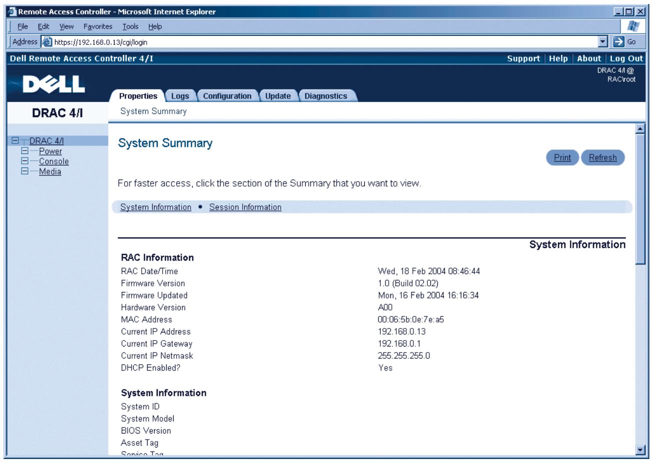
Checking user privileges. When a user logs in from a browser, the DRAC 4 checks the user's privileges during the user authentication phase. After login, the user's privileges determine which tabs and associated subtabs are displayed on the screen. The user authentication process is designed to prevent a user who does not have sufficient privileges from using certain features. For example, a user in the Administrator group can view the tabs and subtabs of all available features after login (see Figure 2). By contrast, a user in the Guest User group is limited to seeing only the tabs and subtabs associated with the Guest User group.

When an administrator uses the remote Racadm utility to manage a system, user authentication and privilege checking are also required. A user who lacks sufficient privileges will fail to execute any command that requires that type of privilege. For example, the configuration command issued by a user without RAC Card Configuration and RAC User Configuration privileges will fail, and the error message will indicate that the user does not have the privilege to execute the command.

When a BMC remote management connection is configured on a server—via serial, LAN, or SOL links—an application or utility that complies with the IPMI 1.5 specification can access the server through that connection.

Racadm utility security

The Racadm utility is a CLI-based tool that can be used to configure and manage the DRAC 4. This scriptable utility can be


Remote Access Controller - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address https://192.168.0.13/cgi/login
Go
Dell Remote Access Controller 4/I
Support | Help | About | Log Out
DRAC 4/I @ RAC\root
DELL
Properties
Logs
Configuration
Update
Diagnostics
System Summary
DRAC 4/I
DRAC 4/I
Power
Console
Media
Print
Refresh
System Information • Session Information

System Summary

For faster access, click the section of the Summary that you want to view.

System Information

RAC Information

RAC Date/Time	Wed, 18 Feb 2004 08:46:44
Firmware Version	1.0 (Build 02.02)
Firmware Updated	Mon, 16 Feb 2004 16:16:34
Hardware Version	A00
MAC Address	00:06:5b:0e:7e:a5
Current IP Address	192.168.0.13
Current IP Gateway	192.168.0.1
Current IP Netmask	255.255.255.0
DHCP Enabled?	Yes

System Information

System ID
System Model
BIOS Version
Asset Tag
Service Tag

Figure 2. Main GUI page for a user in the Administrator group

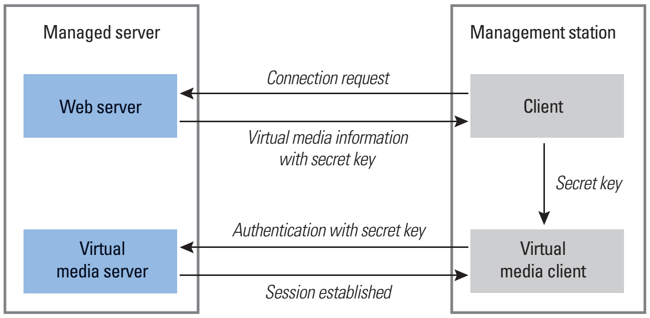


Figure 3. Virtual media security architecture

installed on the managed server or installed on a remote management station client system. The Racadm utility installed on the local managed server is called the local Racadm. The Racadm utility installed on the remote management station is called the remote Racadm.

Local Racadm security. The local Racadm utility communicates with the DRAC 4 through its in-band Peripheral Component Interconnect (PCI) virtual universal asynchronous receiver-transmitter (UART) interface. Because it is installed on the local managed server, administrators need to log in to the managed server to run this utility.

The local Racadm utility enforces security by requiring the user to have either Administrator group or root user privileges to run the utility and access the DRAC 4. On a server running the Microsoft Windows® OS, a user must have Administrator group privileges on the system to run the local Racadm utility; otherwise, an error message is displayed. On a server running the Linux® OS, a user must log in to the system as a root user to have sufficient rights to run the local Racadm utility.

Remote Racadm security. The remote Racadm communicates with the DRAC 4 through its out-of-band NIC. Remote Racadm utility security has been enhanced through the use of an SSL channel to the DRAC 4. A user must successfully pass SSL and user authentication and also have sufficient privileges to execute a Racadm command. Because the remote Racadm utility uses an SSL channel, the commands and data are encrypted by SSL. The RSA (1,024-bit), RC4 (128-bit), and MD5 cipher suite is used in remote Racadm and DRAC 4 SSL communication.

Virtual media security

Virtual media is a powerful remote access feature that provides eighth-generation Dell PowerEdge servers with a virtual CD drive and a virtual floppy disk drive that can use standard media connected anywhere on the network. Administrators can use the virtual media feature from any client on the network to perform various administrative tasks such as OS installation, remote diagnostics, and remote driver and software application installation. To help prevent an attack on a virtual media server, the DRAC 4 uses a security exchange protocol in the virtual media connection.

When a user logs in to the DRAC 4 Web server and selects the Virtual Media subtab, a request-for-connection command is sent to the DRAC 4 firmware (see Figure 3). The DRAC 4 firmware responds by sending virtual media configuration information along with a secret key via a secure SSL channel. Virtual media client software starts a connection and sends its secret key to the virtual media server for authentication. If the secret key passes the virtual media server authentication, a virtual media session is established. Otherwise, a message is sent back to the client indicating that the authentication failed, in which case the connection is dropped. To prevent a replay attack, the secret key is a sufficiently long random number that is dynamically generated by the DRAC 4 firmware each time a request-for-connection command received.

The DRAC 4 enhances current Dell RAC offerings by providing features such as role-based user authentication, Racadm utility security, virtual media security, and console redirection security.

To use the virtual media feature, a user needs the RAC Virtual Media privilege. A user who does not have this privilege will not be able to see the Virtual Media subtab after login.

Dell's virtual media server port number is configurable to help organizations meet their firewall criteria. An administrator can use the Racadm utility to easily configure the port number. The command syntax is as follows:

```
racadm config -g cfgRacVirtual -o cfgVirAtapiSvrPort
   port_number
```

Console redirection security

The DRAC 4 can continuously redirect a managed server's video data to a management station and a management station's keyboard and mouse control to a managed server. The console redirection feature is easy to use and does not require the installation of any special software on either the managed server or the management station. The console redirection feature enables administrators to control a geographically distant server while at a remote management station just as if they were physically present at the managed server.

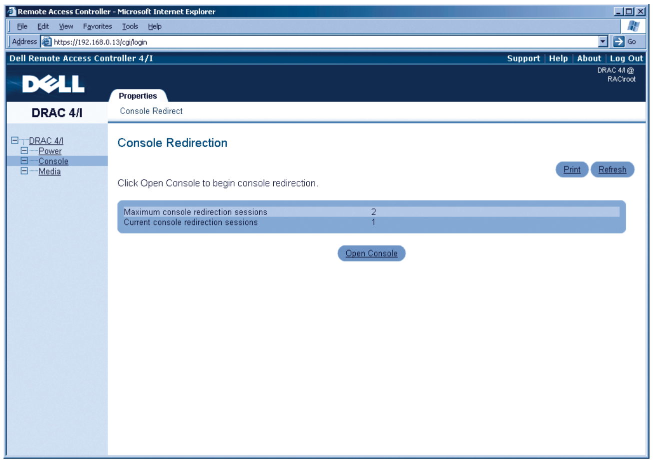
A security protocol has been implemented in the console redirection design to help keep clients that have not been authenticated through the DRAC 4 Web server login from accessing the console redirection path. This design helps prevent a hostile party from interpreting keyboard keystrokes by snooping on the network traffic during remote console redirection.

The following sequence of security protocol operations establishes a console redirection session:

1. The administrator logs in to the main GUI and clicks the Open Console button (see Figure 4). The main GUI sends a preauthentication request to the DRAC 4's embedded Web server via a secure SSL channel (see Figure 5).
2. The DRAC 4 Web server returns secret information including an encryption key via an SSL channel. The console redirection secret information and encryption key are dynamically generated to prevent a replay attack.
3. The console redirection client sends a login command to the console redirection server for authentication. If the authentication is successful, a console redirection session and a console redirection pipe are established. Video, mouse, and keyboard data are redirected in this pipe. Keyboard and mouse data is encrypted on the management station side using an encryption key, and the data is decrypted by the DRAC 4 console redirection server. This makes a network snooping attack virtually impossible.

Enhanced remote management capabilities and security compliance

IT administrators can take advantage of enhanced remote management capabilities by using the on-board BMC on IPMI 1.5–compliant, eighth-generation Dell PowerEdge servers and by properly configuring and maintaining the BMC's remote management connection. At the same time, enhanced DRAC 4 capabilities can further enhance remote server management through informed use of the latest built-in security features.



Click Open Console to begin console redirection.

Maximum console redirection sessions	2
Current console redirection sessions	1

Figure 4. Main GUI Console Redirection page

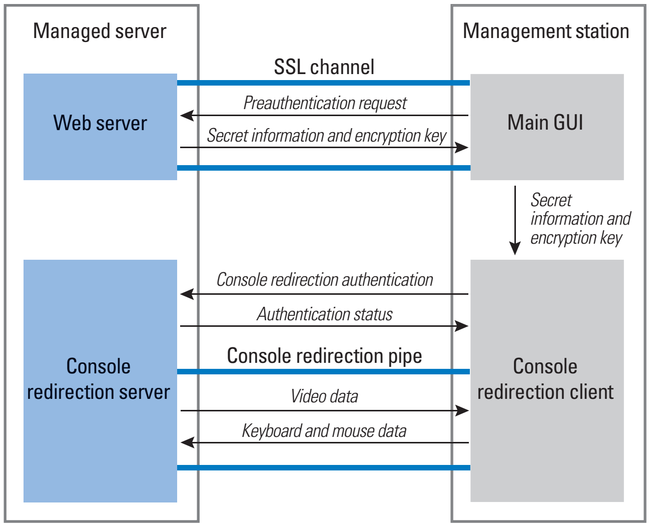

Managed server
Management station
SSL channel
Web server
Preauthentication request
Secret information and encryption key
Main GUI
Secret information and encryption key
Console redirection authentication
Authentication status
Console redirection pipe
Console redirection server
Video data
Keyboard and mouse data
Console redirection client

Figure 5. Console redirection security architecture

Chandra S. Mugunda is a senior development engineer on the Dell Instrumentation Software team. Chandra has an M.S. in Computer Science from the India Institute of Technology, Roorkee, and a B.S. in Electrical Engineering from Andhra University in India.

Weimin Pan is a senior development engineer in the Dell Remote Management Group. He has worked as a senior systems engineer in the Dell Storage Enclosure Subsystem Group. Weimin has an M.S. in Electrical Engineering from the University of Utah and an M.S. in Computer Engineering from Shanghai Jiao Tong University in China.

Haihong Zhuo is a software engineer consultant in the Dell Enterprise Software Development Group. She has worked on systems management solutions and is currently on the Systems Management Instrumentation team. Haihong has an M.S. in Computer Engineering from The University of Texas at Austin and a B.S. in Electrical Engineering from Tshinghua University in China.

FOR MORE INFORMATION

IPMI 1.5 overview and specification:
www.intel.com/design/servers/ipmi/index.htm

Advancing KVM Technology

Through the Dell 180AS Console Switch and 2160AS Console Switch

Administrators must maximize the utility of the available space in the data center while controlling the operational expenses associated with the management of servers and infrastructure. Dell keyboard, video, mouse (KVM) technology can be an integral tool for helping data center administrators achieve these goals. The Dell™ 180AS Console Switch and the Dell 2160AS Console Switch are designed to provide scalable, reliable, and easy-to-use KVM access to multiple servers in heterogeneous environments, making these switches a valuable addition to the data center.

BY MAX A. BENHAM AND ROBERT BERNSTEIN

Related Categories:

Avocent

Data center density

Data center technology

Dell PowerEdge servers

Keyboard video mouse (KVM)

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

In today's data center environments, network administrators depend on keyboard, video, mouse (KVM) switches for cost-effective control of and access to multiple racks of servers. By streamlining the task of accessing servers in dense environments, and by helping facilitate the concentration of IT expertise in a few locations, KVM technology can aid administrators in efficiently managing a large number of servers. However, reliance on KVM technology for 24/7 access to and control of enterprise business servers requires KVM switches on which administrators can depend.

Recently developed by Avocent, the Dell 180AS Console Switch and the Dell 2160AS Console Switch (see Figure 1) are designed to provide the functionality and reliability that organizations demand for access to mission-critical business assets. Developed to suit a variety of business needs, the 8-port 180AS switch allows one administrator to directly access up to 8 servers, while the 16-port 2160AS switch allows two administrators to directly access up to 16 servers. The 180AS and 2160AS switches are designed to maintain reliable connectivity in the data center with predicted mean time between failures (MTBF) of 358,205 hours and 267,165 hours, respectively.[1]

In addition, organizations can use 180AS and 2160AS switches to provide convenient and intuitive accessibility through the Avocent® On-Screen Configuration and Activity Reporting (OSCAR®) graphical user interface (GUI), as well as local console security features, Category 5 (Cat 5) cable connections, cost-effective scalability options, and multiplatform integration capabilities.

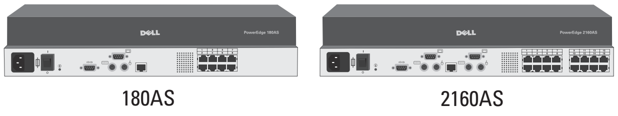


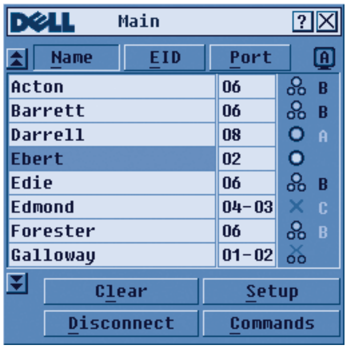
Figure 1. Dell 180AS Console Switch and 2160AS Console Switch

Using the OSCAR GUI for convenient and intuitive accessibility

The 180AS and 2160AS switches use the familiar and intuitive OSCAR GUI developed by Avocent (see

[1] Avocent has calculated these values using the count-part prediction method as defined by Telcordia SR-332, a standard for the reliability prediction of commercial electronic components. The initial values were then modified to reflect expected field results based on Avocent customer research conducted in September 2003 regarding the performance of similar products in the field. The predicted values represent performance at 50 degrees Celsius. For more information about Telcordia SR-332, visit www.isograph-software.com/rwbovertel.htm.

Figure 2). The OSCAR GUI provides easy-to-use, mouse-driven menus for switch configuration, server selection and management, and administrator access.



Acton	06	B
Barrett	06	B
Darrell	08	A
Ebert	02	
Edie	06	B
Edmond	04-03	C
Forester	06	B
Galloway	01-02	

Figure 2. OSCAR overlaid menu system

The OSCAR GUI provides menus both to configure the KVM switch and to select the server that administrators want to control. Through a keyboard hot-key sequence, an administrator can bring up the OSCAR menu at any time, regardless of the OS. By default, pressing the Print Screen key activates the OSCAR GUI; however, administrators can readily configure another activation key sequence.

From the OSCAR GUI, an administrator can select and access any server connected to the switch. The server names are maintained in a list within the switch, and the OSCAR GUI establishes an out-of-band connection to the server, enabling an administrator to access a server without requiring either the network or server OS. The on-screen display provides a flag that indicates which server an administrator is viewing. By accessing the OSCAR GUI and selecting a different server, administrators can switch between servers.

Streamlining console and server management tasks

To streamline typical console and server management tasks, administrators can use the OSCAR GUI to scan servers attached to the switch. The Scan function connects to a server's video signal for an administrator-configurable amount of time and then scans the next server. The administrator can program the list of servers so that only certain servers are scanned, and so that the servers are scanned in a particular order.

Administrators can also use Broadcast mode to streamline typical management tasks. For example, broadcast mode enables the administrator to perform identical tasks simultaneously across multiple servers. Using Broadcast mode, the administrator can perform actions on one server and then select multiple servers on which those actions can be duplicated.

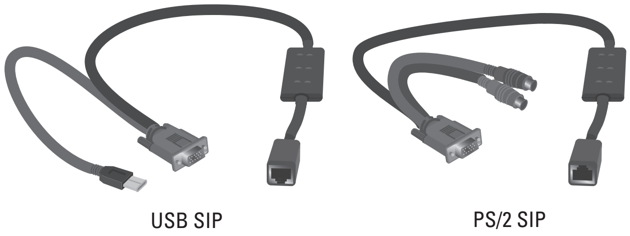


Figure 3. Dell USB SIP and Dell PS/2 SIP

Enabling local console security

The OSCAR GUI has the option of enabling a screen saver with a programmable activation time. The screen saver can be password protected to help prevent access by unauthorized users. Once administrators authenticate with the correct password, they have access to all the connected servers. Both the screen saver and the password protection can be disabled by the administrator.

Benefiting from the use of Cat 5 cables and SIPs

The incorporation of Avocent server interface pod (SIP) technology into the AS product line allows the 180AS and the 2160AS switches to use the same Dell SIPs that administrators already use with the Dell 2161DS Remote Console Switch (see Figure 3), or to use select Avocent SIPs. A SIP drives standard KVM analog signals over a single Cat 5 cable, helping organizations avoid the traditional bulky KVM cables that proliferate in most data centers. By connecting SIPs to servers and then running Cat 5 cables between the RJ-45 ports on the back of the AS switch and each RJ-45 connection on the SIP, administrators can enable five key benefits:

- A reduction in cable density—saving valuable rack space and helping reduce heat within the rack
- The capability to field-terminate Cat 5 cable, easing cable management
- Quick setup and installation of racks and servers because Cat 5 cables are more malleable than the thicker, traditional KVM cabling
- A potential cabling distance of up to 50 feet between an AS switch and a server attached to the switch, allowing for multirack management
- The capability to attach the switch not only to PS/2 servers, but also to USB servers, Sun servers, and serial servers and devices simply by using different SIPs—thereby facilitating management of heterogeneous data center environments

Each type of SIP has built-in memory that stores configuration information for each attached server, enhancing both accessibility and manageability. By linking SIPs to servers, administrators can configure cabling between the switch and the SIP without having to remember which server is attached to which port on the back of the switch. If a cable on the rear of the switch is moved to a different RJ-45 system port, the integrated intelligence in every SIP is designed to make the adjustment within the OSCAR GUI automatically so that the server configuration appears correctly in the interface.

Furthermore, each SIP is powered directly from the attached server to help provide Keep Alive functionality within the data center—meaning that even if the 180AS or 2160AS is not powered up, servers using a SIP have the capability to behave as though a local keyboard and mouse are still attached.

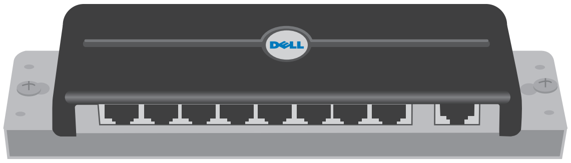


Figure 4. Dell port expansion module

Enabling cost-effective capacity expansion

As business requirements grow and change over time, 180AS and 2160AS switches can help provide the flexibility to accommodate a dynamic data center environment. Administrators can cost-effectively expand the capacity of 180AS and 2160AS switches using the methods that best suit their organizational needs.

Using PEMs to add capacity while minimizing space

By connecting a port expansion module (PEM) to any RJ-45 system port on a 180AS or 2160AS switch, administrators can expand the capacity of a single port from one server to eight servers (see Figure 4). In this way, adding PEM support to 180AS or 2160AS switches can enable a total system capacity of 64 servers for the 180AS and 128 servers for the 2160AS. Moreover, adding PEM support enables organizations to expand cost-effectively because a PEM is available at a fraction of the cost of an additional switch. In addition, a PEM is typically mounted in a 0U configuration, which means that expanding through the use of PEMs can consume less rack space than expanding through the use of additional switches.

Using ACI technology to add capacity and enhance accessibility

Although PEMs provide the capability to economically expand the capacity of either the 180AS switch or the 2160AS switch, in some situations multiple administrators require simultaneous access to servers that are attached to a PEM. In such instances, administrators may prefer to tier a 2160AS switch instead of using a PEM to expand the capacity of the switch. By running a single Cat 5 cable between the Avocent Console Interface (ACI) port of the tiered switch and the RJ-45 port of a primary switch, administrators can use the local console ports of both the primary switch and the tiered switch to gain access to the servers beneath the tiered switch. Tiering 2160AS switches can expand the capacity of the 180AS switch and the 2160AS switch to 128 servers and 256 servers, respectively.

Integrating into an existing KVM over IP environment

The 180AS and 2160AS switches have been designed to seamlessly integrate into a data center's existing KVM over IP infrastructure. When a single Cat 5 cable connects the ACI port of either a 180AS or 2160AS switch to an RJ-45 system port of a Dell 2161DS Remote Console Switch, the local console interface and the Dell Remote Console Software (RCS) Java-based client interface can recognize the servers attached to the tiered switch. In this scenario, additional SIPs are not required to tier the 180AS or 2160AS switch, and the administrator does not have to reconfigure either the tiered AS switch or the 2161DS primary switch.

Controlling dense data center environments

By providing direct access to servers from a single console, 180AS and 2160AS switches help optimize valuable data center space and help eliminate the need for crash carts as well as multiple keyboards, monitors, and mice within server racks. The reliability and usability of these switches can be of critical importance to administrators who require constant access to servers and other devices. Designed to be economical, scalable, flexible, and intuitive, 180AS and 2160AS switches provide functionality that enables administrators to help minimize the cost of managing servers and the server infrastructure while helping to optimize available data center space. By providing support for multiple server platforms and integration with KVM over IP infrastructure, 180AS and 2160AS switches can help provide seamless control over the entire data center.

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.

Robert Bernstein is a product marketing manager for racks and rack peripherals in the Dell Enterprise Systems Group. Previously, he assisted with advanced system sales in the Dell Small and Medium Business Group for eight years. Robert has a B.S. in Communications from The University of Texas at Austin and is a Microsoft Certified Systems Engineer (MCSE).

Avocent Corporation is a leading worldwide supplier of KVM switching, remote access, and serial connectivity solutions that provide IT managers with access to and control of multiple servers and network data center devices.

FOR MORE INFORMATION

Dell 180AS Console Switch and 2160AS Console Switch:
www.dell.com/downloads/global/products/pedge/en/svrac_180as_2160as.pdf

Dell 2161DS Remote Console Switch:
www.dell.com/downloads/global/products/pedge/en/svrac_2161ds.pdf

Exploring the Integrated KVM Capabilities of the Dell PowerEdge 1855 Blade Server

Dell™ PowerEdge™ blade servers, which include a built-in keyboard, video, mouse (KVM) switch, are optimized for the high-density computing environments of today's data center. This article explains the connection options and benefits of integrated KVM switching in the recently released Dell PowerEdge 1855 blade server chassis.

BY STEPHEN M. HAHN AND RYAN FRANKS

Related Categories:

Avocent

Blade servers

Dell PowerEdge servers

Keyboard video mouse (KVM)

Remote management

Systems management

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

Many enterprise data centers are increasing server density in an effort to keep pace with ever-expanding business needs. To maximize the use of costly floor space, many organizations are turning to platforms such as the high-performance, high-density Dell PowerEdge 1855 blade server. Besides the traditional server components required for high-performance environments—including processors, memory, and storage—high-density computing environments require fully functional data center components integrated into a single form factor.

Keyboard, video, mouse switches

To help save space, reduce management costs, and eliminate as much cable clutter as possible, many administrators implement keyboard, video, mouse (KVM) switches. KVM switches provide control for multiple servers and other devices from a keyboard, video monitor, and mouse. Using a KVM switch, administrators can control hundreds of servers from a single console such as the Dell 1U keyboard, video monitor, and mouse tray. The most visible benefit of the KVM approach is the elimination of a physical keyboard, video monitor, and mouse for each server. In addition, managed servers need not be homogeneous. Despite substantial differences, Microsoft® Windows®, UNIX®, and Linux® systems can be controlled from the same KVM console. Both blade servers and KVM switches are optimized for corporate data centers, Internet service providers (ISPs) and application service providers (ASPs), and high-performance computing (HPC) cluster environments where space is often at a premium.

Dell PowerEdge 1855 blade server

The Dell PowerEdge 1855 blade server can be an optimal platform for dense data center environments. Its chassis houses as many as 10 server blades in a 7U form factor and incorporates a standard, built-in Avocent® Analog KVM switch module (see Figure 1). The Avocent Analog KVM module includes a single port for attachment of a KVM dongle, which provides a direct connection to an external keyboard, monitor, and mouse. This approach is designed to provide full control of server functions, including local OS installation, server blade configuration, daily maintenance, and troubleshooting. In addition, the

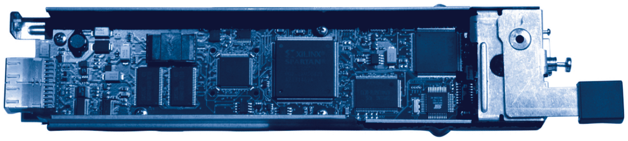


Figure 1. Avocent Analog KVM switch module for the Dell PowerEdge 1855 blade server

KVM switch operates independently of the OS, so individual server blades can run whatever OS is required.

By incorporating a KVM switch into the blade server chassis, organizations can dramatically reduce the amount of cabling and real estate required in the data center. Because of the integrated switching capability, only one set of KVM cables is needed for all 10 server blades—and in a dense data center environment, excessive cabling can become a troublesome issue. The substantial reduction in KVM cables helps the PowerEdge 1855 blade server reduce cable sprawl compared to a traditional 1U server.

An Avocent Analog KVM module can be installed in the Dell PowerEdge 1855 blade server in lieu of a typical analog KVM module. In some blade servers, administrators must connect a bulky KVM dongle to each server blade, and physically switch the dongle from one server blade to another for local console connectivity. Because up to 60 blade servers can occupy a rack, locating the correct server blade and plugging the dongle into it can be extremely inconvenient. In contrast, a PowerEdge 1855 blade server equipped with the Avocent Analog KVM module allows a single Category 5 (Cat 5) cable to provide KVM control for an entire chassis—enabling administrators to access hundreds of server blades from a single, on-screen menu.

Using the combination of the PowerEdge 1855 blade server, the Avocent Analog KVM module, and the Dell 2161DS Remote Console Switch KVM switch, an organization can monitor and control a virtually unlimited number of servers without burying its data center in cables.

KVM connection options

The PowerEdge 1855 blade server provides three options for connecting a KVM console to the server blades using the optional dual-port KVM module:

- Connect the keyboard, video monitor, and mouse cables directly to the PS/2 and video ports on the Avocent Analog KVM module (see Figure 2). The on-screen menu built into the KVM switch lets the administrator select each of the 10 server blades in the chassis.
- Connect the KVM dongle between the PS/2 and video ports on the built-in KVM switch to an external analog or digital KVM switch from Dell or Avocent. Individual server blades for the PowerEdge 1855 blade server then appear in the menu for the external switch just as any directly connected server blade would.
- Connect to an external Dell 2161DS switch through the Analog Console Interface (ACI) port on the built-in KVM switch. This port uses a Cat 5 cable instead of a bulky KVM dongle. The ACI converts the keyboard, video monitor, and mouse signals from the servers, and drives them through a single Cat 5 cable to the external switch. The ACI port is shown in Figure 2.

Integration into the existing infrastructure

The KVM switch on the PowerEdge 1855 blade server can be integrated into a data center's existing KVM infrastructure. When the built-in switch is connected to an external KVM switch such as the Dell 2161DS switch, the interface recognizes the server blades in the chassis and lets the administrator select them from the server list.

In fact, because the PowerEdge 1855 blade server controls all 10 server blades in the chassis, the system consumes fewer external KVM switch ports than traditional servers, which can lead to significant savings in both cost and physical space. Administrators can interact with the server blades in the same way they would interact with stand-alone servers.

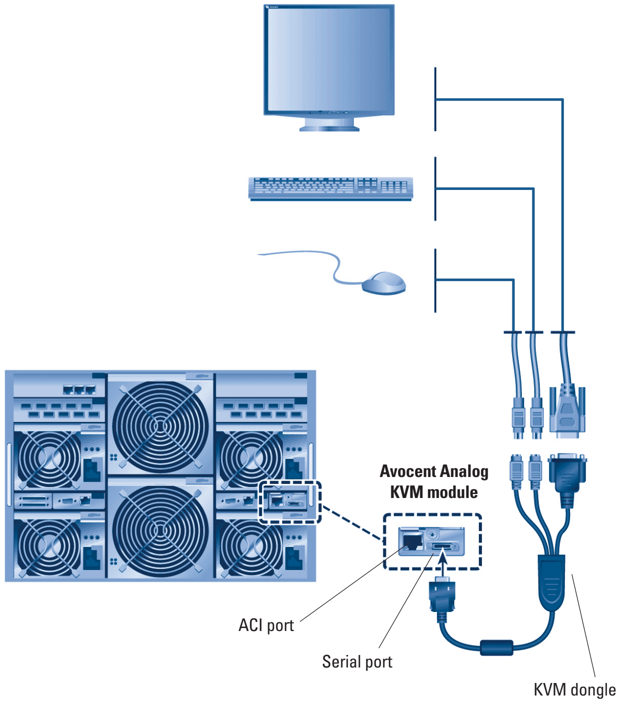
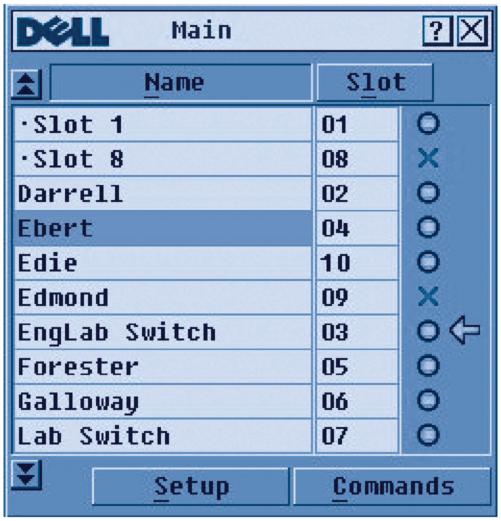


Figure 2. Connecting a KVM console to server blades using the Avocent Analog KVM module



Name	Slot	
·Slot 1	01	○
·Slot 8	08	✕
Darrell	02	○
Ebert	04	○
Edie	10	○
Edmond	09	✕
EngLab Switch	03	○ ⇦
Forester	05	○
Galloway	06	○
Lab Switch	07	○

Figure 3. On-screen menu built into the Avocent Analog KVM module

Use of the management interface

The KVM module of the PowerEdge 1855 blade server uses the Avocent On-Screen Configuration and Activity Reporting (OSCAR®) overlaid menu system. The OSCAR interface provides intuitive menus to configure the KVM switch and select the server that the administrator wants to control through the console (see Figure 3).

OSCAR is a graphical user interface (GUI) used to navigate through connected servers using the mouse and the keyboard. Keyboard hot-key sequences allow the administrator to bring up the OSCAR interface menu at any time, regardless of the OS used. The default sequence is to press the Print Screen key once or the Ctrl key twice; however, additional activation key sequences can be programmed.

From the OSCAR interface menu, the administrator selects a server blade from a list of server blades connected to the built-in KVM switch in the PowerEdge 1855 blade server. The list shows up to 15 characters of the names that the network administrator specified for each server blade when the system was installed. OSCAR can access every server blade connected to the switch. The on-screen display provides a flag that indicates the server blade to which the KVM switch is currently connected.

Securing OSCAR

To avoid leaving the console tuned to one particular server, the OSCAR interface can scan the server blades that it controls. The scan function connects to a system's video signal for a predetermined amount of time, and then moves to the next server blade. The administrator can program both the list of server blades and the amount of time that each server blade is viewed.

The OSCAR interface has the option of enabling a screen saver with a programmable activation time. The screen saver can be password protected to prevent access from unauthorized users. Once a user authenticates with a password, that user has access to all of the server blades in the chassis. To secure individual server blades, the network administrator should enable the OS security for each particular server blade. The screen saver and password protection can be disabled if desired.

Controlling multiple systems

When the KVM switch built into the PowerEdge 1855 blade server is combined with external KVM switches, the OSCAR interface can provide access to hundreds of systems from the same menu. Such access includes all the server blades in the PowerEdge 1855 blade server chassis and any other servers connected through the external KVM switches—even those that are on different racks. The administrator can access any of the connected servers through a single OSCAR interface menu.

Efficient control of data center systems

Data centers for corporations, ISPs, and ASPs are often densely populated with servers and management systems. Blade servers and KVM switches are well suited to such environments, and the PowerEdge 1855 system combines both server blades and integrated KVM switching into a single chassis.

The built-in KVM switch integrates easily into a data center's existing KVM infrastructure. Connectivity is enabled by plugging a keyboard, video monitor, and mouse directly into the switch on the back of the PowerEdge 1855 blade server chassis; connecting to an existing analog KVM switch using the PS/2 and video ports; or plugging the built-in switch into an external Dell 2161DS KVM switch using the ACI port on the built-in switch.

Together, the Dell PowerEdge 1855 blade server hardware and the Avocent OSCAR management software can provide efficient control of data center systems. The PowerEdge 1855 blade server, with its built-in KVM switch, is optimized for the high-density environment of the modern data center.

Stephen M. Hahn has been the Dell worldwide account manager at Avocent for the past four years. Previously, Steve was a technical sales engineer at Avocent.

Ryan Franks is currently a marketing manager in the Dell Enterprise Product Group and has also worked in the Dell Advanced Systems Group. Ryan has a B.B.A. in Marketing from The University of Texas at Austin and is currently pursuing an M.B.A. from The University of Texas at Austin.

Avocent Corporation is a leading worldwide supplier of KVM switching, remote access, and serial connectivity solutions that provide IT managers with access to and control of multiple servers and network data center devices.

Extending Altiris Inventory Solution

with Dell OpenManage for Servers and Clients

To enhance enterprise systems management capabilities, administrators can integrate Dell™ OpenManage™ Server Administrator and Dell OpenManage Client Instrumentation with Altiris® Inventory Solution® software, which is designed to provide consolidated access to comprehensive systems information. Aggregating detailed device status and configuration data into a single view through the Altiris console enables convenient inventorying, reporting, and analysis of Dell hardware, helping organizations to streamline IT management tasks and proactively address security and maintenance issues.

BY TODD MITCHELL AND HECTOR VALENZUELA

Related Categories:

Altiris

Asset management

Change management

Dell OpenManage

Dell OptiPlex desktops

Dell PowerEdge servers

Systems management

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

Tools that quickly and accurately inventory IT assets throughout the enterprise can enhance predictive systems maintenance, expedite the analysis and resolution of security vulnerabilities, and reduce the time and expense associated with systems management. In addition, a consolidated view of enterprise-wide system assets can be instrumental in helping administrators control a distributed IT infrastructure and respond flexibly to ever-changing business requirements.

Altiris Inventory Solution is a powerful software platform that is designed to gather detailed systems information including hardware status, configuration data, installed software packages, and OS settings. Robust inventory capabilities enable Altiris Inventory Solution to scan devices that run a Microsoft® Windows® OS—such as servers, desktops, and notebooks—and collect data from those scanned devices over a network. Data from Altiris scans is stored in a centralized repository, providing administrators with a consolidated view of computing resources across the enterprise.

When used in combination with Dell OpenManage tools, Altiris Inventory Solution offers powerful centralized systems management and reporting capabilities. In particular, Altiris Inventory Solution enables administrators to collect and consolidate the Dell-specific hardware information gathered by Dell OpenManage Server Administrator (OMSA) and Dell OpenManage Client Instrumentation (OMCI) into a single, convenient location. OMSA is a secure Web tool designed to manage and obtain the status of individual Dell PowerEdge™ servers, while OMCI allows administrators to configure and inventory hardware-specific data for Dell desktops and notebooks.

Integrating Dell and Altiris systems management capabilities

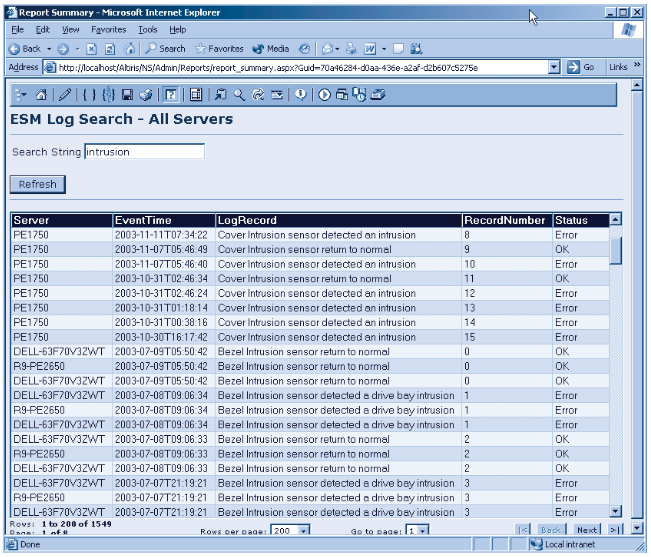
The standard Altiris Inventory Solution scan collects hundreds of application-, OS-, and hardware-related values and stores them centrally in the Altiris database from which administrators can view and analyze the consolidated data. In addition, administrators can generate customizable Web reports from the Altiris console.

Although Altiris Inventory Solution is designed to gather considerable information out-of-the-box, administrators can easily customize the Altiris software platform to collect additional information of interest to their own organizations. This capability to quickly and flexibly extend the standard Altiris inventory scan enables administrators to collect entirely different data sets from a variety of sources, including Windows Management Instrumentation (WMI).

Both OMSA and OMCI publish Dell-specific hardware data values to WMI. Using a simple .xml configuration file, administrators can configure Altiris Inventory Solution to collect the OMSA and OMCI data published to WMI. This integration allows enterprises to collect a rich variety of Dell hardware data and aggregate that data into a single enterprise view to enable easy analysis and troubleshooting of Dell hardware configurations and status.

By leveraging the breadth of information published by OMSA and OMCI, the Altiris platform's centralized management capabilities can help enable administrators to streamline systems management, enhance security, and facilitate predictive maintenance. The consolidation of Dell hardware–specific data allows administrators to run enterprise-wide queries to track system inventory (such as hard drive make and model), monitor security events (such as chassis intrusions), and proactively address maintenance concerns (such as high CPU temperatures) for multiple hardware devices in the environment.

The capability to review data from many devices at once through one centralized resource—the Altiris console—can help make otherwise labor-intensive system administration tasks fast and efficient. For example, OMCI publishes more than 200 Dell hardware data values to WMI. Using Altiris Inventory Solution to collect these data values enables administrators to make detailed assessments about devices in the environment—such as determining which notebooks in the enterprise use a lithium ion battery chemistry—simply by running an Altiris scan and generating a Web report instead of having to manually check each system. By integrating Dell OpenManage tools into Altiris Inventory Solution, administrators can perform in a few minutes systems inventory and maintenance work that used to take days.


Report Summary - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://localhost/Altiris/NS/Admin/Reports/report_summary.aspx?Guid=70a46284-d0aa-436e-a2af-d2b607c5275e
Go
Links

ESM Log Search - All Servers

Search String intrusion

Refresh

Server	EventTime	LogRecord	RecordNumber	Status
PE1750	2003-11-11T07:34:22	Cover Intrusion sensor detected an intrusion	8	Error
PE1750	2003-11-07T05:46:49	Cover Intrusion sensor return to normal	9	OK
PE1750	2003-11-07T05:46:40	Cover Intrusion sensor detected an intrusion	10	Error
PE1750	2003-10-31T02:46:34	Cover Intrusion sensor return to normal	11	OK
PE1750	2003-10-31T02:46:24	Cover Intrusion sensor detected an intrusion	12	Error
PE1750	2003-10-31T01:18:14	Cover Intrusion sensor detected an intrusion	13	Error
PE1750	2003-10-31T00:38:16	Cover Intrusion sensor detected an intrusion	14	Error
PE1750	2003-10-30T16:17:42	Cover Intrusion sensor detected an intrusion	15	Error
DELL-63F70V3ZWT	2003-07-09T05:50:42	Bezel Intrusion sensor return to normal	0	OK
R9-PE2650	2003-07-09T05:50:42	Bezel Intrusion sensor return to normal	0	OK
DELL-63F70V3ZWT	2003-07-09T05:50:42	Bezel Intrusion sensor return to normal	0	OK
DELL-63F70V3ZWT	2003-07-08T09:06:34	Bezel Intrusion sensor detected a drive bay intrusion	1	Error
R9-PE2650	2003-07-08T09:06:34	Bezel Intrusion sensor detected a drive bay intrusion	1	Error
DELL-63F70V3ZWT	2003-07-08T09:06:34	Bezel Intrusion sensor detected a drive bay intrusion	1	Error
DELL-63F70V3ZWT	2003-07-08T09:06:33	Bezel Intrusion sensor return to normal	2	OK
R9-PE2650	2003-07-08T09:06:33	Bezel Intrusion sensor return to normal	2	OK
DELL-63F70V3ZWT	2003-07-08T09:06:33	Bezel Intrusion sensor return to normal	2	OK
DELL-63F70V3ZWT	2003-07-07T21:19:21	Bezel Intrusion sensor detected a drive bay intrusion	3	Error
R9-PE2650	2003-07-07T21:19:21	Bezel Intrusion sensor detected a drive bay intrusion	3	Error
DELL-63F70V3ZWT	2003-07-07T21:19:21	Bezel Intrusion sensor detected a drive bay intrusion	3	Error

Rows: **1 to 200 of 1549**
Page: **1 of 8**
Rows per page: 200
Go to page: 1

Figure 1. Searching all ESM logs simultaneously

Centralized queries also help minimize the need for administrators to manually review and evaluate individual log file contents. For example, Altiris Inventory Solution can automate the collection of critical Embedded Server Management (ESM) log data from OMSA to offer a consolidated view of Dell servers in the environment—providing centralized server management that helps facilitate enhanced levels of security while reducing the time and effort required for server maintenance. For example, administrators can search log entries across every server in the environment to identify servers that have experienced a specific kind of error, such as a chassis intrusion (see Figure 1).

Running the standard Altiris Inventory Solution scan

Administrators have several options for configuring the policies that govern how and when system scans are executed. Altiris Inventory Solution offers the flexibility to run Dell hardware–specific data scans as part of the standard Altiris inventory scan or as separate tasks on separate schedules to suit specific organizational needs.

Altiris inventory scans operate in either of two modes:

- Agent-based, or *policy-driven,* mode
- Agentless, or *zero-footprint,* mode

In zero-footprint mode, a small executable file takes the place of the agent used in policy-driven mode. The executable file can be configured to run and then delete any traces of its existence from the target system's registry and file system after the scan is performed.

Because no agent is resident in zero-footprint mode, this method allows administrators to avoid dedicating resources on the target system to Altiris Inventory Solution except while the scan is being performed. Policy-driven mode offers different advantages, including the capability for a scan to be scheduled and run without network access to the target system—an approach that can be helpful for enterprises with mobile users that are often disconnected from the corporate LAN.

Either scan method can be used for servers or clients, even when collecting custom information such as Dell hardware–specific data from OMSA and OMCI. Both scan methods result in a comprehensive, enterprise-wide hardware and software inventory that is collected into a single database and can be viewed through the centralized Altiris console or customized Web reports.

Extending the scan to collect Dell OMSA and Dell OMCI data

The process of extending the Altiris Inventory Solution scan to collect WMI-published data from OMSA and OMCI is enabled by a simple .xml configuration file. Altiris provides two sample .xml files at no charge. Organizations can test and customize these files to collect WMI-published OMSA and OMCI data using Altiris Inventory Solution in their own environments.[1]

Ultimately, Altiris inventory scans execute by running a file called AeXInvSoln.exe on each target system. The AeXInvSoln.exe executable file references a corresponding .ini file to provide input options that direct the types of scans to run as well as the scan sequence. To enable both policy-driven and zero-footprint scans, .ini and scan files reside in two different locations on Altiris Notification Server™, the platform on which Altiris Inventory Solution is installed:[2]

- **Policy-driven scan:** The default policies that install with Altiris Inventory Solution are directly accessible on Altiris Notification Server at C:\Program Files\Altiris\Notification Server\NSCap\Bin\Win32\X86\Inventory Solution. The .ini file that drives the scan for these files is AeXInvSoln.ini.
- **Zero-footprint scan:** The zero-footprint scan leverages the same files as the policy-driven scan. However, to accommodate distribution, the files are bundled in a 3 MB executable package called AeXWebInvPkg.exe. This package, which contains all the files necessary for an inventory scan to run on a target system, is located on Altiris Notification Server at C:\Program Files\Altiris\Notification Server\NSCap\Bin\Win32\X86. The .ini file that drives the scan for the zero-footprint package is AeXIsHttp.ini.

To enable either type of scan to collect WMI-published OMSA or OMCI data, administrators must extend the scan to reference a custom inventory scan .xml file. *Note:* The procedures to enable OMSA and OMCI scans are virtually identical, but the contents of the .xml files are somewhat different. The .xml file for the OMSA scan points to different WMI classes and properties than the .xml file for the OMCI scan.

Administrators can extend either a policy-driven or zero-footprint scan to collect custom OMSA data using a procedure similar to the following, which demonstrates the use of an Altiris-provided sample .xml file:

1. Download the AltirisOMSAScan.zip file from www.altiris.com/dellzip/altirisomsascan.zip and extract the DellOMSAScan.xml file.[3]
2. Extend policy-driven scanning by copying the DellOMSAScan.xml file to C:\Program Files\Altiris\Notification Server\NSCap\Bin\Win32\X86\Inventory Solution.

   Enable zero-footprint scanning by using the AeXPkgEditor.exe utility to open the AeXWebInvPkg.exe Web package and add the DellOMSAScan.xml file to the package. The AeXPkgEditor.exe utility is located at C:\Program Files\Altiris\Notification Server\NSCap\Bin\Win32\X86.
3. Modify the appropriate .ini file—either AeXInvSoln.ini or AeXIsHttp.ini—to enable a policy-driven or zero-footprint scan by adding the following line of code (shown in bold) to the default file contents:

```
aexauditpls.exe /hidden /output xml
aexmachinv.exe
aexcustinv.exe /in .\DellOMSAScan.xml /out
   DellOMSAScan.nsi
aexcustinv.exe /in .\AeXCustInvStd.cit /out
   AeXCustInvStd.nsi
aexexchpls.exe /hidden /output xml
aexsnplus.exe /output xml
aexnsinvcollector.exe /hidden /o ..\..\..\EvtInbox
```

   The added line of code directs AeXInvSoln.exe to reference the DellOMSAScan.xml file and scan the resources defined there.

Flexible data model to support Dell data classes

The Altiris-provided sample DellOMSAScan.xml file is configured to gather data from three WMI-published classes for OMSA, and the Altiris-provided sample DellOMCIScan.xml file gathers data from 12 WMI-published classes for OMCI. All of these classes are automatically populated with meaningful information upon installation of OMSA or OMCI on a server or client, respectively.

Administrators can easily expand and customize DellOMSAScan.xml and DellOMCIScan.xml to collect the Dell hardware–related data that is most relevant to them.[4] The scan can gather data values from any Dell WMI classes—including classes that are designed to remain empty unless expressly populated by the administrator. Dell Cost of Ownership (Dell_COO), for example, is one such class. Administrators can choose to use this class to capture important information regarding insurance, cost centers, warranties, leasing data, ownership, repairs, and so on.

Seamless extension of database schemas and GUI

Once administrators have extended the Altiris scan using a valid custom .xml inventory file, Altiris Notification Server has all the

[1] Altiris supports the extension of the standard Altiris scan to collect custom inventory data, and can provide general assistance to Altiris customers regarding this concept. However, Altiris does not provide technical support for sample .xml files or any custom files that individual organizations may develop.

[2] Altiris Notification Server is provided by Altiris free of charge.

[3] To enable Altiris Inventory Solution to collect custom OMCI data, download the AltirisOMCIScan.zip file from www.altiris.com/dellzip/altirisomciscan.zip and extract the DellOMCIScan.xml file.

[4] For more information about modifying the DellOMSAScan.xml and DellOMCIScan.xml sample .xml files to scan for the desired class properties, visit www.altiris.com/docs/partners/WP_Extending_OMSA.pdf and www.altiris.com/docs/partners/WP_%20OMCI.pdf.

information it needs to automatically build the database schemas required to store custom Dell data collected by the extended scan. The Altiris database is not extended by the .xml file itself but is instead triggered by incoming custom data when that data is posted by the Altiris agent to Altiris Notification Server for the first time. When Altiris Notification Server detects a data class not previously encountered, the Altiris Data Loader service in Altiris Inventory Solution builds the database tables needed to accommodate this data before importing the data into the Altiris database—all without manual intervention by the administrator.

Similarly, the Altiris console is automatically extended to accommodate the custom data, allowing Altiris Inventory Solution to support Dell hardware–specific data seamlessly. Administrators need not modify the graphical user interface (GUI); Dell hardware–specific data classes automatically appear in the GUI, grouped together with the data gathered by a standard Altiris scan in an intuitive hierarchy (see Figure 2). Placing Dell hardware data into the centralized Altiris environment provides administrators with a powerful tool for accessing and analyzing critical inventory data from a single location.

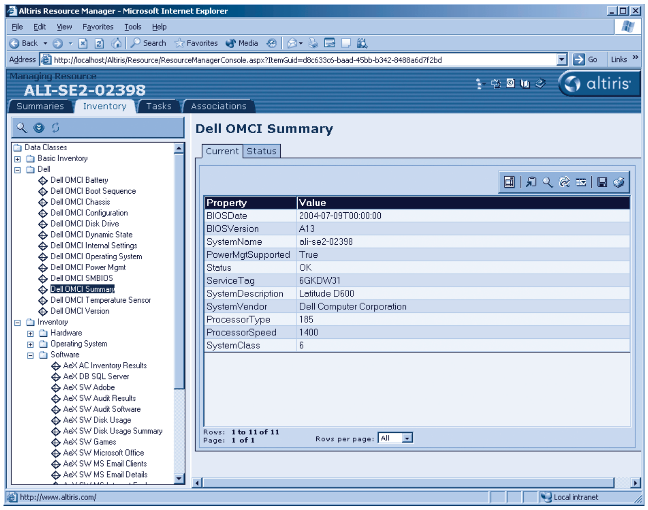
Enabling powerful centralized reporting capabilities

To facilitate viewing and analysis of inventory data, Altiris provides additional sample .xml files that enable three sample Web reports for OMSA data and 36 sample Web reports for OMCI data. *Note:* The procedures to enable OMSA and OMCI Web reports are virtually identical, but the contents of the .xml files are different.

Administrators can install the sample reports for OMSA using a procedure similar to the following, which demonstrates the use of an Altiris-provided sample .xml file:

1. Download the AltirisOMSAScan.zip file from www.altiris.com/dellzip/altirisomsascan.zip and extract the DellOMSAScanReports.xml file.[5]
2. Click the Reports tab in the Altiris console and browse to Assets and Inventory > Inventory > Windows.
3. Right-click on "Windows" and select "Import" from the pop-up menu.
4. Browse to the DellOMSAScanReports.xml file and click "Open." The reports will load into the Altiris console.

Note: Altiris sample reports for OMSA and OMCI were designed to leverage data collected by the sample DellOMSAScan.xml and DellOMCIScan.xml custom inventory scan files, respectively. Without the corresponding database schemas defined in these .xml files, the sample reports will not function properly. To enable the DellOMSAScan.xml or DellOMCIScan.xml file, refer to the procedure in the "Extending the scan to collect Dell OMSA and Dell OMCI data" section in this article.



Dell OMCI Summary

Property	Value
BIOSDate	2004-07-09T00:00:00
BIOSVersion	A13
SystemName	ali-se2-02398
PowerMgtSupported	True
Status	OK
ServiceTag	6GKDW31
SystemDescription	Latitude D600
SystemVendor	Dell Computer Corporation
ProcessorType	185
ProcessorSpeed	1400
SystemClass	6

Rows: 1 to 11 of 11
Page: 1 of 1
Rows per page: All

Figure 2. Integrated Altiris data model

Benefiting from centralized access to hardware information

The capability to collect enterprise-wide hardware status and configuration data from Dell servers, desktops, and notebooks—and to view that detailed systems information from a centralized console—enables administrators to accurately track and manage enterprise-wide IT assets. By integrating Altiris Inventory Solution with Dell OpenManage tools as discussed in this article, administrators can help minimize the time and expense required to collect and analyze critical systems information. In this way, organizations can enhance their capability to successfully implement mission-critical IT infrastructure projects, maintain high standards for systems maintenance and security, and put valuable IT assets to their most effective use—responding flexibly to ever-changing business requirements.

Todd Mitchell is the Dell alliance technical director at Altiris. He has worked with numerous Altiris customers to support Dell-specific implementations and management needs. Todd has a bachelor's degree from Brigham Young University in Provo, Utah.

Hector Valenzuela is a solutions architect in the Custom Solutions Engineering Group at Dell. He has worked on multiple deployment, change management, and storage area network (SAN) projects. Hector has a bachelor's degree in Electrical Engineering from the University of Texas at El Paso.

[5] To install OMCI sample reports, download the AltirisOMCIScan.zip file from www.altiris.com/dellzip/altirisomciscan.zip and extract the DellOMCIScanReports.xml file.

Simplifying IT Operations with Altiris Deployment Solution for Dell Servers

Software tools that enable quick and easy deployment can help IT organizations obtain full server utilization and make the most of limited resources. The Dell™ Deployment Toolkit (DTK) is designed to manage the configuration of hardware in a pre-OS environment, including BIOS and disk configuration. Dell has partnered with Altiris to integrate this hardware configuration capability into Altiris Deployment Solution for Servers, which provides the capability to configure and deploy operating systems and applications to servers. The resulting combination enables administrators to deploy a Dell PowerEdge™ server from bare metal to production with no direct physical interaction beyond the initial hardware setup of plugging in the cables and pressing the power button the first time.

BY TODD MUIRHEAD; DAVE JAFFE, PH.D.; AND LANDON HALE

Related Categories:

Altiris

Dell OpenManage

Dell PowerEdge servers

System deployment

Systems management

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

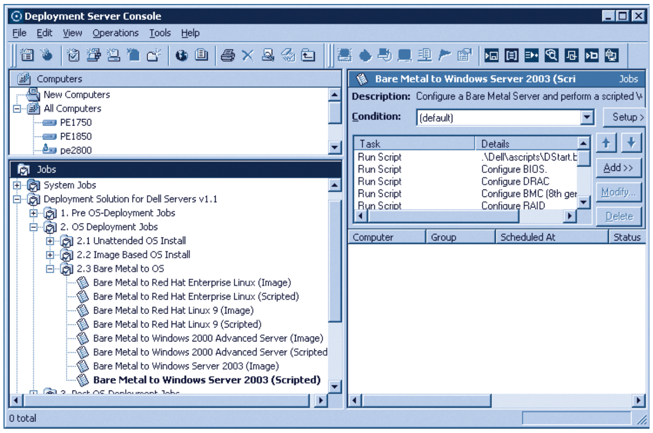
Organizations must be nimble to manage change in complex IT environments that require a rapid business response. Unfortunately, administrators often must grapple with many different tools to manage the overall IT infrastructure, including clients, servers, and other hardware. To help ameliorate this situation, Dell has made simplifying operations a key principle of its scalable enterprise strategy. To that end, Dell works closely with its systems management partners to help streamline deployment, change management, and monitoring capabilities. Dell and Altiris have partnered previously to enhance client management and now they are addressing server management.

Altiris Deployment Solution for Servers offers automation capabilities and drag-and-drop simplicity. The Dell Deployment Toolkit (DTK) provides necessary utilities for managing and configuring Dell server hardware in a pre-OS environment. Altiris has integrated the Dell DTK as an add-on for Altiris Deployment Solution for Servers. The combined package—referred to as the Altiris Deployment Solution for Dell Servers—integrates command-line tools into a set of predefined jobs that can be run from the Altiris console. This approach allows administrators to deploy a Dell server from bare metal to production with no direct physical interaction beyond the initial hardware setup of plugging in cables and pressing the power button the first time.

Altiris Deployment Solution for Dell Servers can address both software and hardware deployment requirements by leveraging the Dell DTK, enabling organizations to deploy and redeploy Dell servers consistently, quickly, and automatically. In fact, customer research conducted by Dell and Altiris in March 2005 indicates that this approach can reduce deployment time from hours to minutes.[1]

[1] "Deployment Comparison for Dell PowerEdge Servers" by Dell and Altiris, www.keylabs.com/results/Altiris/AltirisDeploymentSolution.pdf.

This article focuses on how organizations can leverage Altiris Deployment Solution for Dell Servers to minimize deployment times and enhance deployment consistency by greatly streamlining the process of deployment. The "Deploying the initial system" section in this article discusses how deployment of a Dell PowerEdge server can be accomplished using a script. The "Capturing and deploying the system image and hardware configuration" section shows how the image of the initial server can be used to deploy a second Dell server. The "Automating Dell PowerEdge 1855 server blade deployment" section indicates how Altiris Deployment Solution for Dell Servers can be used to automatically deploy Dell PowerEdge 1855 server blades based on chassis slot location.



New Computers
All Computers
PE1750
PE1850
pe2800

System Jobs
Deployment Solution for Dell Servers v1.1
1. Pre OS-Deployment Jobs
2. OS Deployment Jobs
2.1 Unattended OS Install
2.2 Image Based OS Install
2.3 Bare Metal to OS
Bare Metal to Red Hat Enterprise Linux (Image)
Bare Metal to Red Hat Enterprise Linux (Scripted)
Bare Metal to Red Hat Linux 9 (Image)
Bare Metal to Red Hat Linux 9 (Scripted)
Bare Metal to Windows 2000 Advanced Server (Image)
Bare Metal to Windows 2000 Advanced Server (Scripted
Bare Metal to Windows Server 2003 (Image)
Bare Metal to Windows Server 2003 (Scripted)

Bare Metal to Windows Server 2003 (Scri

Description: Configure a Bare Metal Server and perform a scripted W

Condition: (default)

Task	Details
Run Script	.\Dell\ascripts\DStart.b
Run Script	Configure BIOS.
Run Script	Configure DRAC
Run Script	Configure BMC (8th ger
Run Script	Configure RAID

Computer	Group	Scheduled At	Status

Figure 1. Predefined jobs and subtasks in Altiris Deployment Solution for Dell Servers

Deploying the initial system

To demonstrate the capability of Altiris Deployment Solution for Dell Servers to streamline deployment of Dell servers, a team of Dell engineers set up several Dell PowerEdge servers in a laboratory environment. The team installed Altiris Deployment Solution for Servers 6.1 with Service Pack 1 (SP1) Hotfix D and the Altiris Deployment Solution for Dell Servers add-on on a single-processor Dell PowerEdge server with 512 MB of RAM. The team also installed Microsoft® Windows Server™ 2003, Enterprise Edition; Microsoft SQL Server 2000 with Service Pack 3 (SP3); and Microsoft DHCP Service on the system. *Note:* In production environments, Altiris Deployment Solution should use SQL Server 2000. A Microsoft Data Engine (MSDE) database will also work but is recommended only for test and evaluation environments.

Enabling Microsoft DHCP Service allows systems to automatically obtain addresses during the deployment phase, regardless of whether they will ultimately use Dynamic Host Configuration Protocol (DHCP) or be assigned a static address. Microsoft DHCP Service is not required to run on the same system as Altiris Deployment Solution. A boot floppy can be used instead of DHCP, but requires direct administrator intervention.

Altiris server configuration

To manage servers exclusively over the network using Altiris Deployment Solution for Dell Servers, administrators must ensure that the servers are connected to the same network as the server on which Altiris Deployment Server for Dell Servers is installed. Also, servers should ideally be able to boot via Preboot Execution Environment (PXE); if not, a boot floppy can be used instead of PXE.

To enable PXE on Dell PowerEdge 1855 server blades, PowerEdge 2850 servers, and PowerEdge 6650 servers, the server network interface cards (NICs) must have PXE enabled in the BIOS. By default, PXE is enabled but is not listed first in the device boot order. Administrators must change the server BIOS settings to make PXE the first option in the boot order. This can be accomplished manually by pressing F2 during boot and then making the change directly. Alternatively, administrators can press F12 during initial boot to have the system PXE boot from the Altiris server and then use the Configure BIOS system job—one of many predefined jobs provided in Altiris Deployment Solution for Dell Servers—to change the boot order as part of the initial deployment list of jobs.

For the example deployment described in this article, engineers selected the Simple Install with PXE Server option for the Altiris Deployment Solution for Dell Servers installation. The PXE server included with Altiris Deployment Solution for Dell Servers allows a system with PXE-enabled NICs to boot over the network from the Altiris server. Once the server boots via PXE, the Altiris server can run DOS-based scripts or commands on the system without using local storage or relying on a locally installed OS. This allows for redeployment, imaging, and—using the Dell DTK integrated into the Altiris Deployment Solution for Dell Servers add-on—the capability to reconfigure hardware settings.

Running a job or task on a server from the Altiris console is designed to be a simple drag-and-drop process. Administrators simply click on a job to select it, and then drag and drop the job to the target system. A dialog box presents the administrator with the options of running the job immediately or scheduling it to run at a later time.

Scripted OS deployment

The Dell team installed Windows Server 2003 on a PowerEdge 2850 server using the Bare Metal to Windows Server 2003 (Scripted) job that is provided in Altiris Deployment Solution for Dell Servers (see Figure 1). To prepare for this installation, Dell engineers ran the Get All Configuration Files (BIOS, DRAC, BMC) job against a reference PowerEdge 2850 server to create a set of configuration files based on a PowerEdge 2850 server. The job automatically placed these configuration files in C:\Program Files\Altiris\eXpress\Deployment Server\Dell\Toolkit\Systems\pe2850.

The configuration files act as a template and can be modified if administrators want to change the hardware configuration of future PowerEdge 2850 deployments. The Bare Metal to Windows Server 2003 (Scripted) job automatically uses the configuration files in the pe2850 directory when running on a PowerEdge 2850. Similar directories reside in this location for other Dell PowerEdge server models supported by the Dell DTK including the PowerEdge 1655, PowerEdge 1750, PowerEdge 1800, PowerEdge 1850, PowerEdge 2650, PowerEdge 2800, and PowerEdge 6650 servers as well as the PowerEdge 1855 blade server.

During setup, the script-based installation of Windows Server 2003 on the PowerEdge 2850 used the i386 directory from the Windows Server 2003 installation CD. The entire installation was completed over the network with no local intervention required on the system. The Bare Metal to Windows Server 2003 (Scripted) job also installed the Altiris Deployment Solution client on the PowerEdge 2850 to prepare the server for future deployment jobs.

To complete a system installation, administrators can also use Altiris Deployment Solution for Dell Servers to install additional services. For example, Dell engineers used the Install Dell OpenManage Server Administrator job to load the Dell OpenManage™ infrastructure on the PowerEdge 2850 in the example deployment described in this article. After this step was accomplished, the server was ready for imaging—thus allowing administrators to quickly and easily redeploy its base configuration to other systems.

Capturing and deploying the system image and hardware configuration

Capturing and replicating a system image across multiple servers is very straightforward using Altiris Deployment Solution for Dell Servers. The add-on also allows administrators to configure and replicate system hardware settings. Using predefined jobs, administrators can configure BIOS, Dell Remote Access Controller (DRAC), baseboard management controller (BMC), and PowerEdge RAID controller settings using drag-and-drop operations—enabling administrators to use a single deployment job to make hardware configuration changes to servers across the enterprise. Within the predefined jobs provided in Altiris Deployment Solution for Dell Servers, a single job includes the set of tasks necessary for configuring hardware settings and deploying images.

Using Altiris Deployment Server for Dell Servers, the Dell team took a system image from the PowerEdge 2850 server on which Windows Server 2003 had been installed using the scripted OS installation described earlier in this article. To ensure a complete system configuration, Dell engineers also captured the hardware configuration. The team then installed the hardware configuration and OS image onto a second PowerEdge 2850 as described in the following sections.

Capturing the OS image

Administrators can use Altiris Deployment Solution for Dell Servers to take an image of the system by simply right-clicking on the icon for the system and selecting "Quick Disk Image." Doing so creates a job to capture the OS image and prompts the administrator to schedule the image-capture job. If the image or images must be stored in a location other than the default location, administrators must specify the nondefault location when creating the job.

For example, the server that was used in the example deployment discussed in this article had a C: drive partition of only 10 GB, so the engineers moved the image location to the D: drive partition, which was mapped to a Dell/EMC CX700 storage array on a storage area network (SAN). The Dell team edited the autoexec.bat file of the PXE boot environment to map an I: drive that would allow access to the Dell/EMC CX700 storage array during imaging by adding the following line (shown here in bold) just after the existing line that maps the eXpress share to the F: drive:

```
A:\net\net use F: "\\W2K3ALTIRIS1\eXpress" /yes
A:\net\net use I: \\W2K3ALTIRIS1\images /yes
```

Image capture of the Windows Server 2003, Enterprise Edition, OS with no additional applications installed took Dell engineers about 20 minutes for the example described in this article. The image-capture job accomplished this by shutting down the server and booting into an Altiris-managed environment in which the imaging tool ran and captured the image to the location specified in the job.

Capturing a hardware configuration

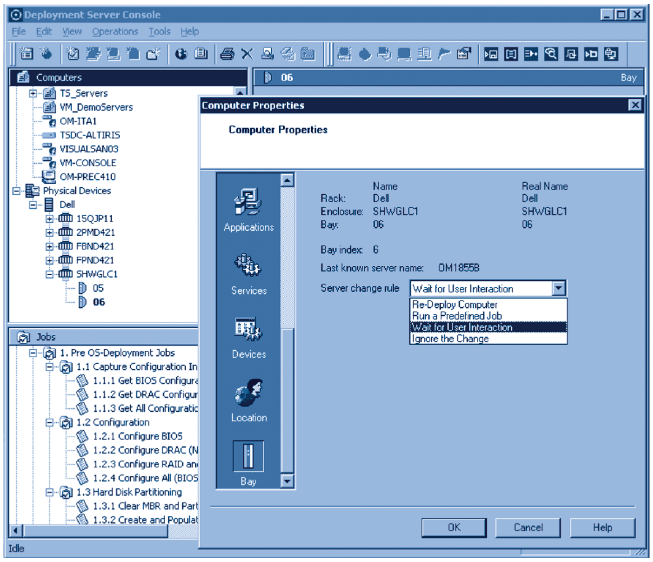
The PowerEdge 2850 server that was initially installed with Windows Server 2003 using Altiris Deployment Solution for Dell Servers had the Altiris Deployment Solution client installed. To capture the BIOS, DRAC, and BMC configuration, Dell engineers ran the Get All Configuration Files (BIOS, DRAC, BMC) job from the Altiris Deployment Solution for Dell Servers on the PowerEdge 2850 server.

The server rebooted several times while the job was running. At the conclusion of the job, the configuration was stored in the C:\Program Files\Altiris\eXpress\Deployment Server\Dell\Toolkit\Systems\pe2850 directory under the BIOS and DRAC subdirectories. The configurations reside in text files that can be edited with Microsoft Notepad or another text editor so that administrators can make changes in the hardware configuration that was captured from the initial system.

Deploying the captured hardware configuration and OS image

To deploy a captured image to another server, Altiris Deployment Solution for Dell Servers provides predefined jobs for deploying a server from bare metal—via scripting or images—for Red Hat® Enterprise Linux®, Windows® 2000 Advanced Server, and Windows Server 2003 operating systems.

To deploy the captured PowerEdge 2850 OS image and hardware configuration onto another PowerEdge 2850, the Dell team used the Bare Metal to Windows Server 2003 (Image) job. Because the images had been captured and stored on the D: drive partition



Computer Properties

	Name	Real Name
Rack:	Dell	Dell
Enclosure:	SHWGLC1	SHWGLC1
Bay:	06	06

Bay index: 6

Last known server name: OM1855B

Server change rule: Wait for User Interaction

Figure 2. Configuring jobs based on server blade slot location

mapped to the SAN, the team was required to change the image path from `F:\images` to `I:\images`. To make this modification, Dell engineers simply selected the image job, scrolled down the list of tasks, clicked "Deploy Task," and then clicked "Modify" to change the path in the image job from `F:\images` to `I:\images`.

Altiris Deployment Solution for Dell Servers lets administrators deploy a captured image to a target server in two ways:

- By predefining the deployment job based on the target server's Media Access Control (MAC) address, Dell service tag, or asset tag
- By waiting for the target system to appear in the Altiris console (which occurs automatically when a system PXE boots from the Altiris server), and then dragging and dropping the job onto the target server

Both methods require that the system PXE boot from the Altiris server or boot from a properly configured boot floppy before the deployment job will run. Deployment jobs include tasks that first configure the hardware settings, then copy the image onto the system, and finally reboot the server to that image. Following the initial boot, the Altiris server can be configured to perform an initial configuration that provides the target system with a unique address, host name, and other settings specific to that system.

Automating Dell PowerEdge 1855 server blade deployment

The Dell PowerEdge 1855 blade server chassis supports 10 individual server blades, offering a compelling density advantage over standalone servers in space-constrained data centers. Besides the imaging and scripting functionality that Altiris Deployment Solution for Dell Servers enables on supported Dell PowerEdge servers, the add-on supports special capabilities specific to PowerEdge 1855 blade servers.

For example, when using Altiris Deployment Solution for Dell Servers, administrators can assign a deployment job based on the slot that a server blade occupies in the PowerEdge 1855 blade server chassis. The capability to recognize which slot a server blade occupies can help automate server blade deployment.

Administrators can configure Altiris Deployment Solution for Dell Servers to automatically perform a variety of actions when a server blade is inserted into a slot (see Figure 2). Options include redeploying the system, running a predefined job, waiting for user interaction, or ignoring the insertion of the server blade into the slot.

The capability to run slot-specific jobs enables what is commonly referred to as rip-and-replace functionality. For example, a specific slot in a PowerEdge 1855 blade server might contain a server blade that is being used as a Web server within a load-balanced farm of Web servers. This server blade could be quickly removed upon failure and a replacement server blade added to the same slot. Administrators could configure Altiris Deployment Solution for Dell Servers to automatically deploy the image for the Web server onto the replacement server blade upon insertion of the server blade into the chassis, helping to save time and minimize administrative efforts.

Enhancing administrator control over server resources

Altiris Deployment Solution for Dell Servers is designed to streamline the deployment of Dell servers. Enabling administrators to install a server from bare metal using the drag-and-drop capabilities of the Altiris interface can greatly simplify systems management. Additionally, the rip-and-replace functionality enabled for Dell PowerEdge 1855 blade servers helps allow for automatic redeployment of systems, especially in load-balanced server farms—enabling convenient, comprehensive control over critical systems resources.

Todd Muirhead is a senior engineering consultant on the Dell Global Alliances team. He specializes in SANs, virtualization, and database systems. Todd has a B.A. in Computer Science from the University of North Texas and is Microsoft Certified Systems Engineer + Internet (MCSE+I) certified.

Dave Jaffe, Ph.D., is a senior consultant on the Dell Global Alliances team who specializes in cross-platform solutions. Previously, he worked in the Dell Server Performance Lab, where he led the team responsible for Transaction Processing Performance Council (TPC) benchmarks and the Dell Technology Showcase. He has a Ph.D. in Chemistry from the University of California, San Diego, and a B.S. in Chemistry from Yale University.

Landon Hale manages Dell's relationship with Altiris within the Dell Global Alliances team. Previously, he worked in various product marketing, sales, and sales management roles at Dell and at Sea-Land Service. Landon has a B.A. in Political Science from Carleton College and an M.B.A. from the Marshall School of Business at the University of Southern California.

Implementing Fault Tolerance

Through VMware Scripting and Dell OpenManage

The VMware® Virtual Infrastructure™ SDK package enables scripts written in popular programming languages to access servers and manipulate the virtual machines (VMs) they host under VMware ESX Server™ software when these hosts are managed by VMware VirtualCenter management software. In this study, the Virtual Infrastructure SDK scripting facility was combined with Dell™ OpenManage™ software to demonstrate a powerful fault-tolerance capability. This article explains how Dell engineers developed a script that called the VMware VMotion™ feature to move all the VMs from a failing server to other servers in a server farm. The failing server was then taken down, serviced, brought back online, and repopulated with the VMs it had been hosting—all with little impact on end-user applications.

BY DAVE JAFFE, PH.D., AND TODD MUIRHEAD

Related Categories:

Dell OpenManage

Dell PowerEdge servers

Fault tolerance

Scripting

Virtualization

VMware

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

For server consolidation, fault tolerance, and ease of administration, many enterprises have moved all or part of their data center applications onto virtual machines (VMs) that reside on multiple Dell PowerEdge™ servers running VMware ESX Server software. Administrators can use VirtualCenter—the VMware central administration program—to manage ESX Server–based servers as a logical pool of resources in a server farm. VirtualCenter lists the physical servers and VMs in the server farm and displays a console showing current information regarding performance, tasks, and events occurring on those servers. VirtualCenter also allows administrators to create, clone, start, and stop VMs, as well as to move live VMs between physical servers running ESX Server software through the use of VMware VMotion technology.[1]

VMware provides the VMware Virtual Infrastructure SDK (software developer's kit) scripting facility, which enables developers to access VirtualCenter data and call VirtualCenter commands through a Web services interface that is accessible using most popular programming languages, including C# and Java.[2] By enabling developers to access VirtualCenter features through a Web services interface, the VMware Virtual Infrastructure scripting facility extends the functionality of VirtualCenter and enables

[1] These features are discussed in "Introducing VMware ESX Server, VirtualCenter, and VMotion on Dell PowerEdge Servers" by Dave Jaffe, Ph.D.; Todd Muirhead; and Felipe Payet; in *Dell Power Solutions,* March 2004.

[2] To download the VMware Virtual Infrastructure SDK package, visit www.vmware.com/support/developer/vc-sdk.

developers to create custom VMware applications that assist them in managing virtual server infrastructures. For example, developers can create a script to clone additional VMs from existing "golden" VM images in a repository, and then tailor each VM as required—for instance, by changing the host name and IP address.

Server administration tools such as Dell OpenManage IT Assistant (ITA) and Dell OpenManage Server Administrator (OMSA) provide other important components needed to manage large server farms running VMs. Administrators can use ITA (the centralized management console) and OMSA (the agent running on each server) to detect and report error conditions, inventory hardware assets, and update server software and firmware. A key feature of ITA is its capability to call administrator-written scripts in response to specific events from servers, such as warnings. This functionality can be combined with VMware scripting capabilities to provide tools that enhance the power and flexibility of a virtual computing infrastructure.

This article describes an example of how the Dell OpenManage infrastructure can be used together with the VMware Virtual Infrastructure SDK scripting facility to load balance a farm of ESX Server–based servers. In December 2004, a Dell test team developed a script to call the VMware VMotion feature to move all the VMs from one ESX Server–based server to another without shutting down the VMs. The script was designed such that, when OMSA detected an error on a server in the server farm, the failing server triggered OMSA to send a warning to ITA, which in turn called the VMware script on the server that ran VirtualCenter administration software. This script, a C# program, used the VMware Virtual Infrastructure script to call the VMware VMotion facility to move VMs from the failing server while the VMs were running. As a result, the VM migration had virtually no impact on end-user applications. Once removed from the server farm, the failed server could be serviced, brought back online, and repopulated with the VMs that had been temporarily moved to other servers.

Using the VMware Virtual Infrastructure scripting facility

The VMware Virtual Infrastructure SDK package enables developers to access VirtualCenter data and call VirtualCenter commands through a Web services interface. Programs communicate with VirtualCenter using the standard Simple Object Access Protocol (SOAP)/XML Web services protocol to access data (from hosts, VMs, performance counters, events, and so on) or to issue commands (such as `Clone`, `Migrate`, and `Power On/Off`). The interface to the VMware VirtualCenter Web service is defined by `vma.wsdl`, the VMware application programming interface (API) Web Services Description Language (WSDL) file. Programming languages that can issue SOAP/XML Web service requests can be used for the client programs. The SDK supplies examples in C#, Java, Perl, and Visual Basic.

Programming client applications in C# is straightforward when following the examples provided in the SDK. The details of the Web service interface are contained in a C# source file called `vmaService_proxy.cs`, which must be compiled with the client source code. In the sample C# files included in the SDK and in the program described in this section of this article, the calls to this proxy have a wrapper of client-callable programs inside the client application.

The program used in this example—VMotion All VMs, or *vmall*—illustrates the use of the VMware Virtual Infrastructure to move the VMs on a given server running ESX Server sequentially to other ESX Server–based hosts in the farm in a load-balanced manner. The program uses the VirtualCenter Web service to list the VMs on the specified host, then check the CPU utilization on the other ESX Server–based hosts in a specified migration pool, and finally, issue the VMotion `Migrate` command repeatedly to move the VMs off the specified ESX Server–based server.

The vmall program's source file is called `vmall.cs`. (To view the `vmall.cs` source file, visit *Dell Power Solutions* online at www.dell.com/powersolutions.) As seen in `vmall.cs`, the vmall program follows the architecture used in the SDK sample files. `VmaClient`, a class defined in the source file, includes VMware code to create instances of the class, plus code to connect to the Web service as well as code to put a wrapper around `vmaService_proxy.cs` functions that manipulate VirtualCenter objects (such as `ResolvePath`, `GetContents`, and `MigrateVM`). The administrator-written function `get_host_cpu_time` is a special function that uses the performance objects exposed by VirtualCenter to return the total CPU time (in milliseconds) used during the previous minute by a server running ESX Server. To increase the accuracy of this sample, administrators can create a performance counter in the VirtualCenter user interface called "Past Hour," which provides one sample per minute for 60 minutes.

The overall vmall functionality occurs in the administrator-written function `Main`. Four steps are required to move all VMs off a server:

1. Connect to the VirtualCenter Web service.
2. Check that the specified source server is part of the migration pool of servers running ESX Server.
3. Determine the number of VMs on the source server running ESX Server, and list the other servers running ESX Server in the migration pool.
4. Iterate through the list of VMs and move each VM to the least-loaded server in the migration pool at the time.

Connecting with the VirtualCenter Web service requires the administrator's credentials (username and password) along with the URL of the Web service. For demonstration purposes, the username and password are hard-coded into the program in plain text and the URL references a non–Secure Sockets Layer (SSL) version of the Web page. In a production scenario, the username and password would be encrypted and the Web service would be accessed via SSL.

The vmall program first creates an instance of the `VmaClient` object (called `vma`), and then uses the URL, username, and password to call the `Connect` method of `VmaClient` (which puts a wrapper around the `vmaService_proxy.cs Login` function) to connect to the VirtualCenter Web service.

Next, vmall checks that the source host (the host from which the VMs will be moved) is a member of the migration pool of ESX Server–based servers. For demonstration purposes, the set of hosts that are available for VM migrations is hard-coded into the program in the array `migration_pool_hostnames`. A production version of this script could determine this set at runtime by querying VirtualCenter for a list of hosts meeting administrator-specified criteria.

The vmall program then uses the VirtualCenter Web service to count the VMs on the source server and list the other ESX Server–based hosts in the migration pool. The method `ResolvePath("/host")` is used to obtain a handle to the VirtualCenter host list, which is used by the `GetContents` VirtualCenter object to capture this list in a `Container` object. The program then iterates through this list of ESX Server–based hosts. If a listed host matches the source host, the number of VMs currently on the source host is determined. If a listed host is not the source but is contained in the `migration_pool_hostnames` array, the number of CPUs on that host is stored in the `hosts_to_migrate_to_n_cpus` array, and the vmall program generates an additional array of host names called `hosts_to_migrate_to`.

Finally, vmall generates a list of VMs on the source host by drilling down into the VirtualCenter host object, and then iterates through this list using the `MigrateVM` object to move each VM to one of the hosts in the `hosts_to_migrate_to` array. The particular host is chosen by checking CPU utilization and identifying the host with the least-loaded CPUs. This is done by using the `get_host_cpu_time` method described earlier in this section to obtain the total CPU utilization time of that host for the last minute (in milliseconds). That value is then converted to a utilization percentage by dividing by the number of milliseconds in a minute (60,000) and by the number of CPUs in the host and multiplying by 100. Each call to `MigrateVM` returns a handle to a task that is used to monitor the progress of the VMotion migrations. The VMotion migrations occur sequentially in vmall, but other methods are possible. The vmall program writes the progress of each migration to the Microsoft® Windows® command prompt from which vmall is called. Administrators can also track the progress of the VMotion migrations by viewing the Tasks tab in VirtualCenter. When the migrations are complete, the program disconnects from the Web service and returns control to the command prompt.

Setting up Dell OpenManage to work with VMware scripts

Dell OpenManage infrastructure comprises a set of systems management tools for monitoring, updating, and managing Dell hardware. Specifically, the ITA centralized console and the OMSA server monitoring and management tool are used for monitoring servers. ITA is typically installed on a single server that acts as the monitoring console. OMSA is installed on each Dell PowerEdge server to monitor that server. OMSA uses Simple Network Management Protocol (SNMP) to send information to ITA about hardware status changes such as a disk failure or an overheated processor. Once the information reaches the ITA console, an administrator-defined set of rules dictates what actions should be taken. For example, if a disk fails on the database server, ITA can be set up to send an e-mail to the database administrator, and possibly run a script to shut down the database or back it up.

Every Dell PowerEdge server (except for PowerEdge SC servers) ships with Dell OpenManage software, which includes OMSA and ITA. Dell supplies versions of OMSA for Windows, Linux®, and Novell® NetWare® operating systems. Servers running VMware ESX Server use the Linux version of OMSA. Because OMSA is designed to monitor the hardware of the ESX Server–based server, VMs running on ESX Server–based servers do not require their own hardware-monitoring capability.

To set up OMSA to run on an ESX Server–based server:

1. Run `/usr/bin/omasetup.sh` from the ESX Server console command prompt. When prompted, insert the ESX Server CD and then the OMSA for Linux CD.
2. Add this line to the end of the `/etc/snmp/snmpd.conf` file:

```
trapsink IP address of IT Assistant server public
```

3. Run the `service snmpd restart` command.
4. Reboot the server.

The ITA central console can be set up to respond to various types of events, either from specified servers or from all servers. A demonstration can be triggered by intentionally mis-setting a fan speed warning value so that OMSA sends a warning that a fan speed is too slow. This demonstration uses an Event Filter that is triggered by environmental warnings (which include fan speed) that come from ESX Server–based servers. To enable this demonstration in the ITA console, select the Configuration dialog box, the Event Filters dialog box, and then the Add dialog box. In the Add dialog box, configure the following settings, and then click the OK button:

- **Filter Name:** ESXDEMO1
- **Severity Configuration:** Select warning only
- **Select Event Categories/Types:** Select all environmental
- **Select Source Nodes:** Select ESX1, ESX2, and ESX3

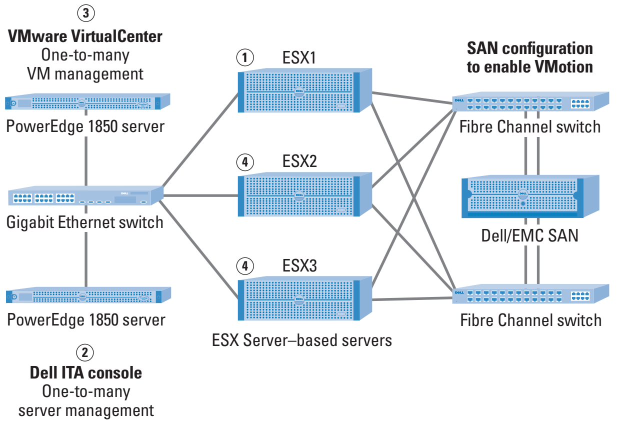


Figure 1. Using Dell OpenManage and VMware software for fault tolerance

This Event Filter also contains a list of actions that will occur when the event is received. In addition to displaying an alert message on the ITA console and writing the event in the Windows event log, the Event Filter creates an action that calls a Visual Basic script called `remote2.vbs`. The `remote2.vbs` script remotely calls the `vmall.exe` program on the server running VirtualCenter and passes it the host name of the ESX Server–based server sending the warning. (To view the `remote2.vbs` script, visit *Dell Power Solutions* online at www.dell.com/powersolutions.) *Note:* The action's executable file is `cscript.exe`—the Windows Script Host (WSH) executable—and the `$n` in the argument list will be replaced at the time the action is called by the host name of the server originating the event. The script `remote2.vbs` passes along this host name to the vmall program as its only argument.

Combining Dell OpenManage capabilities with VMware scripting

The process used by Dell engineers to integrate Dell OpenManage and VMware software to enable automatic migration of VMs from a failing ESX Server–based server is illustrated in Figure 1.

In the configuration demonstrated by Dell engineers, a server farm consisting of three ESX Server–based servers was managed by VirtualCenter running on a PowerEdge 1850 server. A second PowerEdge 1850 server ran ITA to manage the ESX Server–based servers in the data center (not shown in Figure 1). The two PowerEdge 1850 servers were connected to the three ESX Server–based servers and the rest of the data center through a Gigabit Ethernet switch. These three servers were in turn connected to a storage area network (SAN) using redundant host bus adapters (HBAs) in each server, redundant Fibre Channel switches, and redundant connections to the storage.

In the test scenario, OMSA running on ESX1 detected a low fan speed on that server and sent an event to ITA via SNMP. Using the ESXDEMO1 event filter and associated action described in the "Setting up Dell OpenManage to work with VMware scripts" section in this article, ITA reacted to the warning event by calling a local script, `remote2.vbs`, which in turn called the VMware script `vmall.exe` on the VirtualCenter system, passing to it the host name of the system sending the warning. The vmall program then sequentially moved the VMs from the failing ESX Server–based server to other ESX Server–based servers, moving each VM to the least-loaded ESX Server–based server at the time. When this process was completed, the failing server, ESX1, could be taken offline for repairs.

Using scripts to automate tasks and enhance availability

The scenario described in this article demonstrates one of many possible tools that administrators can develop by using the Dell OpenManage infrastructure together with the VMware Virtual Infrastructure SDK scripting facility. The script that the Dell test team developed helped automate fault-tolerance capabilities in a virtual computing infrastructure, and could be used in a real-world scenario to facilitate systems management and enhance application availability for end users. The VMware Virtual Infrastructure SDK package enables administrators to access the features of the VMware VirtualCenter administration program, thereby allowing IT organizations to develop scripts that can automate a wide variety of systems management tasks in addition to the script explored in this article.

Acknowledgments

The authors would like to thank Sudhir Shetty for the Windows Management Instrumentation (WMI) scripts, Mark Clifton for valuable discussions regarding the integration of Dell OpenManage and VMware, and Shivraj Ahluwalia at VMware for support on the VMware scripting API.

Dave Jaffe, Ph.D., is a senior consultant on the Dell Global Alliances team who specializes in cross-platform solutions. Previously, he worked in the Dell Server Performance Lab, where he led the team responsible for Transaction Processing Performance Council (TPC) benchmarks and the Dell Technology Showcase. He has a Ph.D. in Chemistry from the University of California, San Diego, and a B.S. in Chemistry from Yale University.

Todd Muirhead is a senior engineering consultant on the Dell Global Alliances team. He specializes in SANs, virtualization, and database systems. Todd has a B.A. in Computer Science from the University of North Texas and is Microsoft Certified Systems Engineer + Internet (MCSE+I) certified.

FOR MORE INFORMATION

VMware Virtual Infrastructure SDK package:
www.vmware.com/support/developer/vc-sdk

Troubleshooting the Dell PowerEdge 1855 Blade Server

The Dell™ PowerEdge™ 1855 blade server provides several interfaces and utilities that enable administrators to troubleshoot fault conditions in the system chassis and individual server blades, including the Dell Remote Access Controller/Modular Chassis (DRAC/MC) and the baseboard management controller (BMC). This article discusses how to use the DRAC/MC, the BMC, and the Dell OpenManage™ infrastructure along with other hardware and software tools and utilities to help diagnose system issues on the Dell PowerEdge 1855 blade server.

BY MICHAEL BRUNDRIDGE AND RYAN PUTMAN

Related Categories:

Baseboard management controller (BMC)

Blade servers

Dell OpenManage

Dell PowerEdge servers

Out-of-band (OOB) management

Remote management

Systems management

Troubleshooting

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

The Dell PowerEdge 1855 blade server is designed to provide several ways to manage both the system chassis and the individual server blades inside the chassis. These approaches include the Dell Remote Access Controller/Modular Chassis (DRAC/MC) and the baseboard management controller (BMC). The DRAC/MC[1] is accessible using an out-of-band serial port, a Telnet connection, or a Web-based graphical user interface (GUI). The BMC in each server blade is connected to the server blade's integrated network interface card (NIC), which allows administrators to access the BMC using either an Intelligent Platform Management Interface (IPMI) management utility such as IPMI Shell (ipmish) or serial over LAN (SOL) commands.

The Dell PowerEdge 1855 blade server also incorporates several hardware and software components that facilitate troubleshooting and predictive failure notification, including LEDs (see Figure 1). In addition, the system can be managed using the Dell OpenManage suite. Figure 2 summarizes Dell PowerEdge 1855 blade server troubleshooting interfaces and utilities.

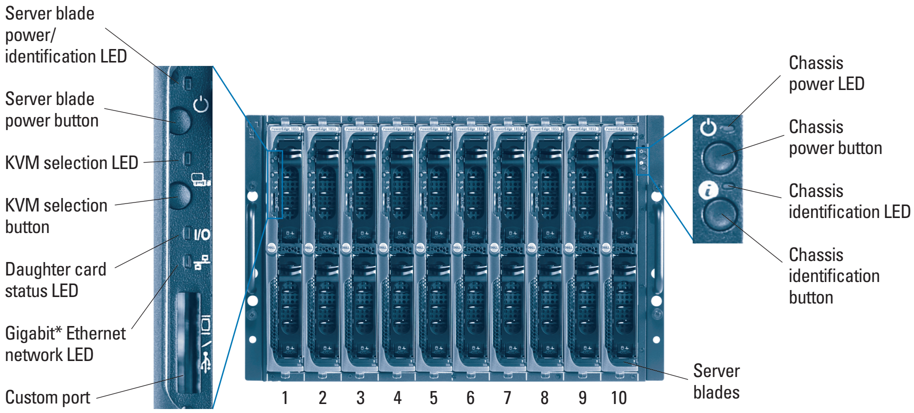
Understanding Dell PowerEdge 1855 blade server logs

The Dell PowerEdge 1855 blade server chassis includes three log file features that can give administrators insight into the system status. Understanding the different roles these log files play is key to effective troubleshooting of the PowerEdge 1855 blade server.

Two of the logs are accessible only through the DRAC/MC: the remote access controller (RAC) log and the chassis system event log (SEL). These two logs identify chassis-level errors only. The RAC log indicates the general status of the system, including errors, while the SEL provides information relating to chassis-level sensor failures.

The third log is the BMC SEL, which indicates failures in individual server blades. Each server blade contains one BMC SEL. The BMC SEL can be accessed in three

[1] The Dell PowerEdge 1855 blade server currently supports only one DRAC/MC. However, redundant DRAC/MC components are planned for a future release.


Server blade power/ identification LED
Server blade power button
KVM selection LED
KVM selection button
Daughter card status LED
Gigabit* Ethernet network LED
Custom port
I/O
1
2
3
4
5
6
7
8
9
10
Server blades
Chassis power LED
Chassis power button
Chassis identification LED
Chassis identification button

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

Figure 1. Dell PowerEdge 1855 blade server chassis and LEDs

ways: through ipmish commands to the BMC, through SOL commands to the BMC, or through Dell OpenManage tools installed on the server blade. *Note:* The BMC SEL cannot be accessed from the DRAC/MC.

Understanding Dell PowerEdge 1855 blade server LEDs

Most components in the PowerEdge 1855 blade server have LEDs that indicate faults (see Figure 3). LEDs are also used to identify components and convey information such as power status. Some LEDs have more than one purpose.[2]

Troubleshooting the chassis

Several tools can be used to troubleshoot chassis-level conditions. The DRAC/MC is the primary tool for troubleshooting chassis faults, whereas the BMC is the primary tool for troubleshooting server blade issues. Fault LEDs provide administrators that have physical access to the system with a visual indication when failures occur.

The first step in troubleshooting the blade server is to determine if AC power is being supplied to the chassis. Each power supply installed in the blade server has a set of three LEDs located at the rear of the chassis. The right-most green indicator on the power supply module is the AC power LED. If this LED is illuminated, AC power is being supplied from the power grid to the chassis. If this indicator is not illuminated for one or more of the power supply modules, administrators should check the AC cabling, power distribution unit, and breakers.

If physical access to the chassis is not convenient, administrators can verify whether the AC power is on by logging in to the DRAC/MC using either the Racadm command-line interface (CLI) or the GUI. (See the "Using the DRAC/MC sensors to troubleshoot shared components" section in the online portion of this article by visiting *Dell Power Solutions* online at www.dell.com/powersolutions.) Because the DRAC/MC operates on standby power, the DRAC/MC is available as long as AC power is present and the power supply modules are functional.

Next, administrators should determine whether the power supply is providing DC power. Administrators can power up the chassis using either the chassis power button or the DRAC/MC. Once the chassis power is on, the power supply modules will supply DC power to the system. The left-most green indicator on the power supply module is the DC power indicator. If this indicator is illuminated, DC power is being supplied from this power supply module to the DC power bus of the chassis. The DRAC/MC can be used to determine power supply status as well. (See the "Using the DRAC/MC sensors to troubleshoot shared components" section in the online portion of this article by visiting *Dell Power Solutions* online at www.dell.com/powersolutions.)

If administrators can successfully log in to the DRAC/MC, the fault can be isolated to a specific server blade or shared component, such as a fan module or an I/O module. If a shared component has a fault, the chassis identification LED will blink once per second. When this occurs, administrators should look for a shared component that also has its fault LED illuminated. To check the current status of the shared components using the DRAC/MC Racadm CLI, administrators can log in to the DRAC/MC console and enter the following Racadm command:

```
racadm getmodinfo
```

The output will resemble the code shown in Figure 4.

Troubleshooting tool	Purpose
LED	Fault indication for both the chassis and individual server blades
DRAC/MC	Out-of-band chassis management
DRAC/MC connect feature	Out-of-band console redirection from a server blade console or an I/O module switch console to the management station
ipmish	Server blade BMC LAN channel access
SOL	In-band console redirection from a server blade console to the management station using the BMC
OMSA	Server blade management

Figure 2. Dell PowerEdge 1855 blade server troubleshooting interfaces and utilities

[2] For a complete list of LED functions and their descriptions, refer to the *Dell PowerEdge 1855 Systems Installation and Troubleshooting Guide* at support.dell.com/support/edocs/systems/pe1855.

Troubleshooting Dell PowerEdge 1855 server blades

The integrated BMC in each server blade is designed to be a service processor. The BMC is responsible for maintaining the status of the server blade and providing administrators with the capability to remotely manage the server blade.

Performing basic power checks

System administrators who have convenient access to the front of the server blade can determine basic server blade status from the LED display. If the server blade is receiving power (but currently turned off), then the server blade power LED will be solid amber and the NIC LED will be solid green. If an administrator selects the server blade with the KVM button, even if the server blade is powered down, the KVM LED will turn solid green.

If the server blade is not receiving power, administrators should try reseating the server blade or placing it in a different slot. *Note:* If a problem exists in the disk subsystem, the LEDs on the server blade's disk drive have a different implementation than the LEDs on other Dell servers.[3]

```
#<group> <module>   <presence> <pwrState> <health> <svcTag>
 1  -->  Chassis    Present    ON         OK       GLCR001
 1  -->  Fan-1      Present    ON         OK       GLCR001
 1  -->  Fan-2      Present    ON         OK       GLCR001
 1  -->  PS-1       Present    ON         OK       GLCR001
 1  -->  PS-2       Present    ON         OK       GLCR001
 1  -->  PS-3       Present    ON         OK       GLCR001
 1  -->  PS-4       Present    ON         OK       GLCR001
 2  -->  DRAC/MC-1  Present    ON         OK       GLCR001
 3  -->  DRAC/MC-2  Absent     N/A        N/A
 4  -->  Switch-1   Present    ON         OK       GLCR001
 5  -->  Switch-2   Present    ON         OK       GLCR001
 6  -->  Switch-3   Absent     N/A        N/A
 7  -->  Switch-4   Absent     N/A        N/A
 8  -->  Server-1   Present    ON         OK       GLB04
 9  -->  Server-2   Present    ON         OK
10  -->  Server-3   Absent     N/A        N/A
11  -->  Server-4   Present    ON         OK       GLB17
```

Figure 4. `racadm getmodinfo` command output

To troubleshoot the server blade using in-band applications and methods, the BMC must be present and functional. Once the server blade has power, administrators can access the BMC, which opens up a variety of options for continued troubleshooting. The most appropriate troubleshooting method will depend on the architecture of the management network.

Component	LED	Description	LED state
Chassis	Chassis power LED	Fault (power off)	Amber blinking fast (twice per second)
		Fault (power on)	Green blinking fast (twice per second)
	Chassis identification LED	Identification	Amber blinking slow (once per second)
Server blade	Server blade power/ identification LED	Fault (power off)	Amber blinking fast (twice per second)
		Fault (power on)	Green blinking fast (twice per second)
		Identification (power off)	Amber blinking slow (once per second)
		Identification (power on)	Green blinking slow (once per second)
	Keyboard, video, mouse (KVM) selection LED	KVM selected	Solid green if selected
		Fault (server module not selected)	Amber blinking
		Fault (server module selected with the KVM)	Green/amber blinking
	Daughter card status LED	Firmware error exists	Green blinking fast (twice per second)
	Gigabit Ethernet network LED	No link from the server blade to the Gigabit Ethernet switch or pass-through module	Off

Figure 3. Dell PowerEdge 1855 blade server LED descriptions

Using Dell OpenManage to troubleshoot server blades

If a server blade can load its OS, then the Dell OpenManage suite provides an alternative troubleshooting approach. Using Dell OpenManage, a system administrator can ascertain the server blade's general health status, view the server blade's SEL, and obtain the current status of the server blade's various sensors. Dell OpenManage also provides access to Dell OpenManage Storage Services (OMSS), which allows system administrators to manage the disk subsystem. *Note:* After power-on self-test (POST), OMSS is the only way to obtain the status of the server blade's disk subsystem.[4]

Using Dell OpenManage Server Administrator (OMSA) requires a user login account with the appropriate privileges. The OMSA console can be run either remotely or locally through a Web browser session. Dell OpenManage IT Assistant (ITA) can be used for remote monitoring of the modular chassis, its shared components, and the server blades contained in the chassis. ITA also has the capability to launch remote instances of the OMSA console for server blades (provided that OMSA is installed on the server blade) as well as a remote instance of the DRAC/MC Web console.

Using ipmish to troubleshoot server blades

The ipmish utility resides on the Dell OpenManage Server Assistant CD. This out-of-band utility allows system administrators to manage the BMC over the network. Besides Dell OpenManage, the ipmish utility provides the main method of viewing the server blade's SEL.

[3] For a complete list of LED functions and their descriptions, refer to the *Dell PowerEdge 1855 Systems Installation and Troubleshooting Guide* at support.dell.com/support/edocs/systems/pe1855.

[4] For more information about troubleshooting the disk subsystem on the Dell PowerEdge 1855 blade server, visit *Dell Power Solutions* online at www.dell.com/powersolutions to view the online portion of this article.

Before administrators use ipmish, they must first assign the BMC an IP address. To do this, administrators can enter the BMC setup program during POST. The BMC setup screen also allows administrators to create users and assign passwords.

To obtain the status of a server blade, administrators can enter the following ipmish command at an OS command prompt:

```
ipmish -ip bmc_ip_address | bmc_hostname
    -u username -p password sysinfo
```

One of the fields in the output generated by the preceding command indicates the general health of the server blade. This command also returns version information.

To obtain the last 10 entries in the server blade's SEL, administrators can enter the following ipmish command at an OS command prompt:

```
ipmish -ip bmc_ip_address | bmc_hostname
    -u username -p password sel get -last 10
```

If a fault condition is hardware related, or if the BMC is notified of the problem, the SEL should contain an error message regarding the fault.[5]

Using SOL to troubleshoot server blades

The BMC on each server blade supports SOL. This feature requires the SOL Proxy agent to be running on a management station or other system located on the network. Once enabled, the primary function of SOL is to provide console redirection from the server blade through the network to the management station. This console redirection connection does not depend upon the DRAC/MC.

Note: Administrators must select the server blade's console redirection configuration in the system's BIOS before OS boot. The console redirection port can be set to BMC SOL, DRAC/MC, or Off. If set to DRAC/MC or Off, the console redirection port cannot be used by SOL until the system is reset, at which time the BMC can be reconfigured to use the console redirection port.

Using the DRAC/MC to troubleshoot server blades

The primary troubleshooting feature available in the DRAC/MC is console redirection from the server blade to the DRAC/MC console. After enabling console redirection in the BIOS prior to OS boot by selecting the DRAC/MC, administrators can log in to the DRAC/MC console and enable console redirection by entering the following command from the DRAC/MC CLI:

```
connect server-x
```

where *x* is the server blade location, or *slot.*

Using console redirection to troubleshoot server blades

It is important to note that a given server blade can support only one console redirection at a time, either through the DRAC/MC or through the BMC using SOL. Console redirection can be a powerful tool for troubleshooting fault conditions because it allows remote administrators to view everything that occurs during POST and to interact with the BIOS, RAID controller, and BMC—all without traveling to the server blade's physical location. Furthermore, administrators can boot to a USB device connected to the front dongle and interact with the DOS prompt. For example, administrators can boot to a USB key that contains Dell 32-Bit Diagnostics and MP Memory diagnostic tools. Administrators can remotely run these tools and observe their results.

In addition to BIOS interaction through console redirection, both the Microsoft® Windows Server™ 2003 and Linux® operating systems support access to the OS via a serial port. In Windows Server 2003, administrators can access the Special Administration Console (SAC). In Linux, administrators can access a text console. Both approaches allow administrators to troubleshoot the OS.

Enhancing the troubleshooting process

Despite the PowerEdge 1855 blade server's increased rack density and shared components, administrators can approach troubleshooting and ongoing support functions in much the same way as they do for stand-alone servers. The PowerEdge 1855 blade server's integrated local and remote management capabilities are designed to help administrators efficiently perform remote troubleshooting and management activities. For best-practices information about troubleshooting the PowerEdge 1855 blade server's shared components, I/O modules, and disk subsystem as well as other troubleshooting factors, visit *Dell Power Solutions* online at www.dell.com/powersolutions.

Michael Brundridge is a technologist in the Dell Enterprise Software Development Group. Before joining Dell, he worked as a hardware engineer for Unisys. Michael attended Texas State Technical College and holds a technical degree from Southwest School of Electronics.

Ryan Putman is a platform developer for the Dell Enterprise Product Group. He has a bachelor's degree in Electrical Engineering from Vanderbilt University and a master's degree in Computer Engineering from North Carolina State University.

FOR MORE INFORMATION

Dell PowerEdge 1855 Systems Installation and Troubleshooting Guide:
support.dell.com/support/edocs/systems/pe1855

[5] For definitions of error messages and additional ipmish commands, see the *Dell OpenManage Baseboard Management Controller User's Guide* at support.dell.com/support/edocs/software/smbmcmu.

MEGABYTE:
What not having a Linux strategy can take out of your bottom line.

If you're paying unreasonable licensing fees for software that constantly needs security patches, you're getting eaten alive. But there's a solution. With SUSE® LINUX, Novell® can help you unleash the cost-saving power of a flexible, end-to-end open source strategy. Only Novell supports Linux from desktop to server, across multiple platforms. We'll integrate our industry-leading security, management and collaboration tools seamlessly into your environment. We'll provide award-winning technical support 24/7/365, and train your IT staff to deploy Linux-based solutions. And we'll make sure your open source strategy actually meets your number-one business objective – making money. Call 1-800-215-2600 to put some teeth back into your tech strategy, or visit www.novell.com/linux ➔ **WE SPEAK YOUR LANGUAGE.**

suse
A NOVELL BUSINESS

Novell®