DELL POWER SOLUTIONS · FEBRUARY 2007

POWER SOLUTIONS

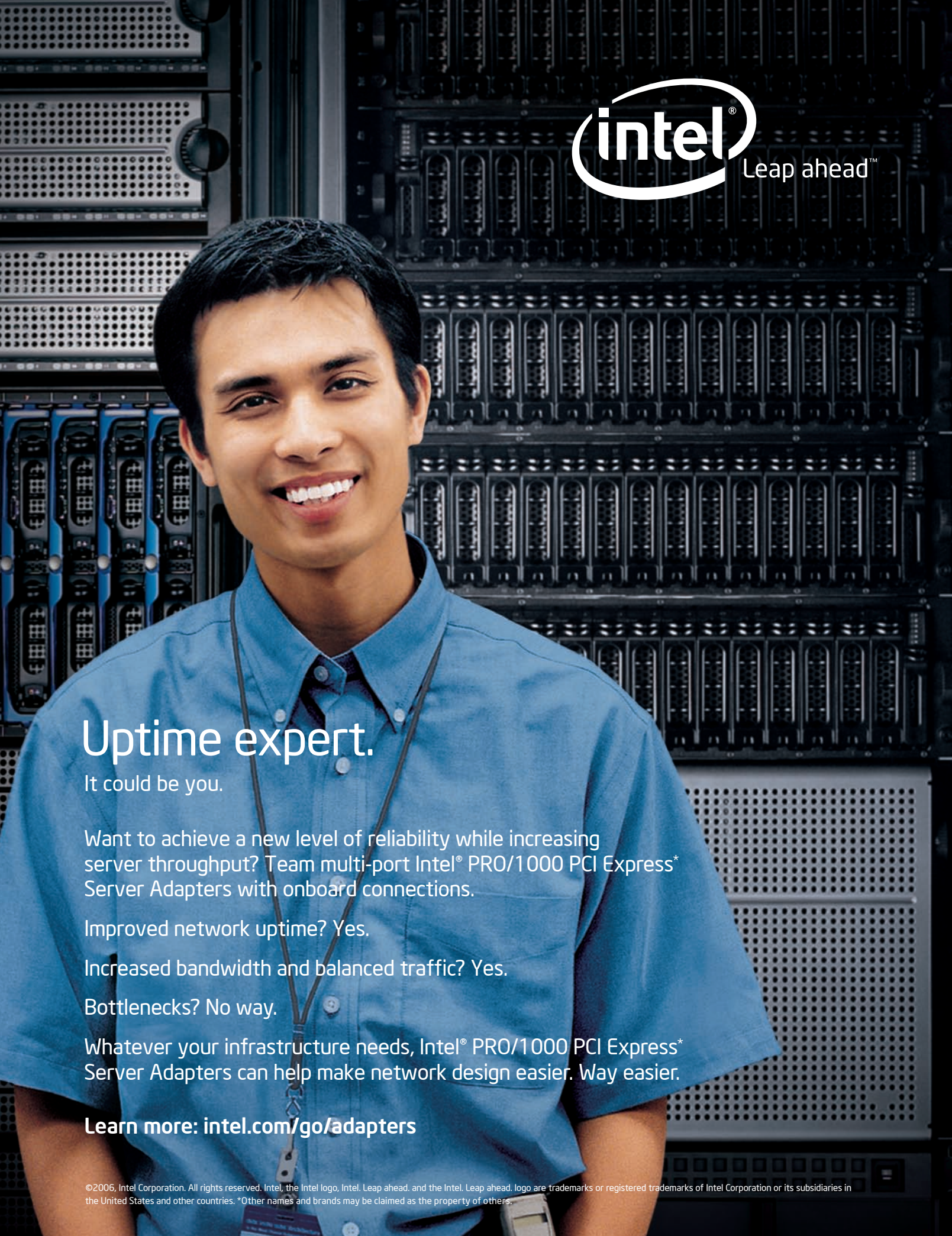
TRANSFORMING YOUR IT FRAMEWORK INTO A SCALABLE ENTERPRISE

FEBRUARY 2007 • $4.95

Powering and Cooling the Data Center

Energy-Efficient Systems Help Lower TCO

Powering and Cooling the Data Center: Energy-Efficient Systems Help Lower TCO


intel
Leap ahead™
Uptime expert.
It could be you.
Want to achieve a new level of reliability while increasing server throughput? Team multi-port Intel® PRO/1000 PCI Express* Server Adapters with onboard connections.
Improved network uptime? Yes.
Increased bandwidth and balanced traffic? Yes.
Bottlenecks? No way.
Whatever your infrastructure needs, Intel® PRO/1000 PCI Express* Server Adapters can help make network design easier. Way easier.
Learn more: intel.com/go/adapters
©2006, Intel Corporation. All rights reserved. Intel, the Intel logo, Intel. Leap ahead. and the Intel. Leap ahead. logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. *Other names and brands may be claimed as the property of others.

POWER SOLUTIONS

TRANSFORMING YOUR IT FRAMEWORK INTO A SCALABLE ENTERPRISE

February 2007

COVER STORY | PAGE 8

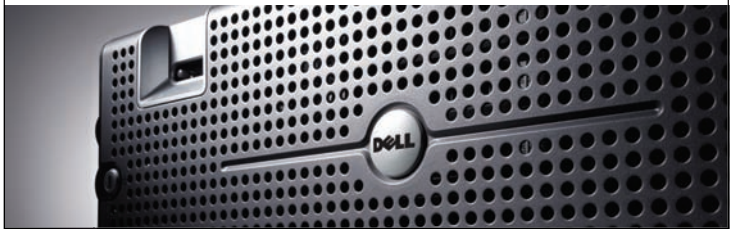
Data Center Efficiency in the Scalable Enterprise

By John Pflueger, Ph.D., and Sharon Hanson

Energy efficiency is a top concern as IT managers struggle with rising power bills, cooling problems, usage limits imposed by local utilities, or requirements to deploy additional servers without expanding an existing data center or building a new one. To address energy efficiency effectively, data center managers must look at a variety of issues ranging in scope from the smallest piece of silicon to the entire data center.

STEP BY STEP | PAGE 22

Efficient Power Management on Dell PowerEdge Servers with AMD Opteron Processors
By Manoj Gujarathi, Lisa Garrison, and Jordan Hargrave
IT administrators can manage power efficiently by enabling AMD PowerNow! technology on AMD Opteron processor–based PowerEdge SC1435 and PowerEdge 6950 servers.



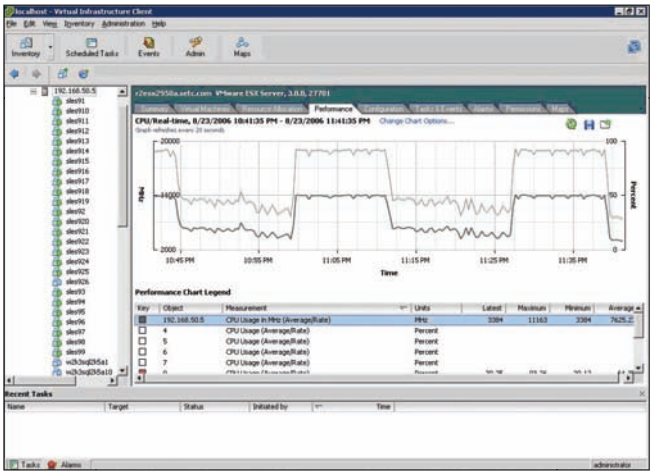
The power-friendly Dell PowerEdge 6950 server provides outstanding performance per watt for demanding enterprise applications. See page 24.

Dell OpenManage Newsletter

Join us in a one-on-one dialogue with the people who are architecting Dell's management strategy for the scalable enterprise. This issue focuses on Dell and Altiris tools that offer a comprehensive approach to client management throughout the enterprise.

DELL ENTERPRISE TECHNOLOGY CENTER SERIES

STORAGE



Monitoring processor utilization with the VMware Virtual Infrastructure 3 virtualization suite. See page 28.

PROFESSIONAL WORKSTATIONS

DELL ENTERPRISE SOFTWARE UPDATES

The Dell PowerVault NX1950 is a unified network storage solution that simultaneously stores both file and application data and supports a wide range of operating environments and communication protocols, including Internet SCSI (iSCSI). See page 43.

DELL POWER SOLUTIONS TRANSFORMING YOUR IT FRAMEWORK INTO A SCALABLE ENTERPRISE

WWW.DELL.COM/POWERSOLUTIONS

SUBSCRIPTIONS

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

ABOUT DELL

Dell Inc., headquartered in Round Rock, Texas, near Austin, listens to customers and delivers innovative technology and services they trust and value. Uniquely enabled by its direct business model, Dell sells more systems globally than any other computer company, placing it 25th on the Fortune 500 list. For more information, visit our Web site at **www.dell.com**. To get Dell news direct, visit **www.dell.com/rss**.

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

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

February 2007

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**.

Now Playing: Dell Datacenter of the Future Podcast Series

In this podcast series, Dell and industry experts discuss key enterprise infrastructure strategies designed to grow your business. The series will cover a broad range of areas including power and cooling, electronic messaging, virtualization, high-performance computing clusters, databases, collaboration, backup and recovery, and other important emerging technologies. For more information, visit **www.dell.com/podcast.**

ADVERTISER INDEX

ONLINE EXTRAS

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. Plus: These *Dell Power Solutions* articles are available only online at **www.dell.com/powersolutions**.

Configuring Dell OpenManage IT Assistant to Monitor Non-Dell Servers

By Saiprasad S.

Dell OpenManage IT Assistant 8.0 introduces the Simple Network Management Protocol (SNMP) Event Source Import utility, which allows administrators to import multiple event source definitions from non-Dell servers.

Configuring Favorite Application Launch Points in Dell OpenManage IT Assistant

By Manoj Poonia

Dell OpenManage IT Assistant 8.0 enables administrators to configure launch points for management applications and tools for discovered networked devices.

Introducing Dell SAS RAID Storage Manager for Dell PowerEdge SC Servers and Dell Precision Workstations

By Manjusha Gopakumar and Gagan Shrestha

The Dell Serial Attached SCSI (SAS) RAID Storage Manager application enables administrators to configure, monitor, and manage storage configurations on SAS controllers, physical drives, and virtual disks.

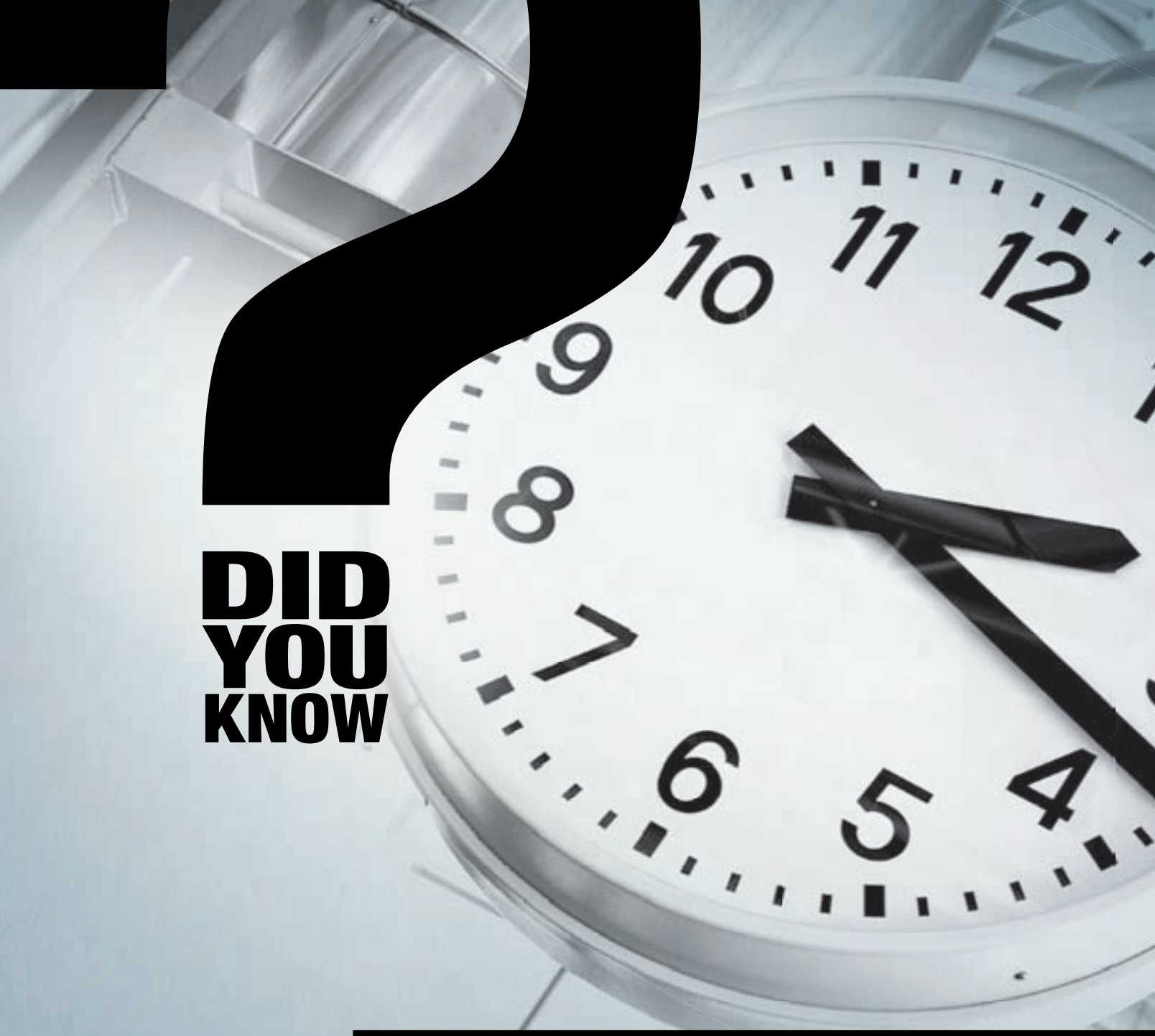
Creating Fault-Tolerant Xen Virtualization at the Network Adapter Layer

By Scott Collier and Brian Gautreau

Using the Linux bonding driver with standardized networking configurations enables highly available, fault-tolerant physical layers for Xen guest domains.

Your Information, Your Way

Can't wait? For real-time access to content as soon as it's posted, subscribe your favorite RSS (really simple syndication) reader to the feed for *Dell Power Solutions* online at **www.dell.com/powersolutions** or **www.dell.com/rss**. Searching for specific article content? Visit our Related Categories index online at **www.dell.com/powersolutions**.

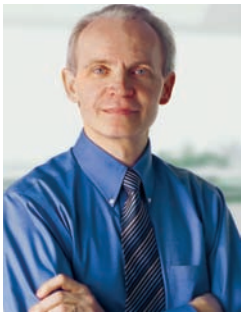


DELL NEVER SLEEPS SO YOU CAN!

Dell's Enterprise Command Centers can give customers high-level support 24/7 for server and storage environments. From minor issues to full-scale national disasters, Dell Enterprise Command Centers have the know-how to get your critical systems up and running again. There's much more to Dell than you know – find out at **www.dell.co.uk/didyouknow**

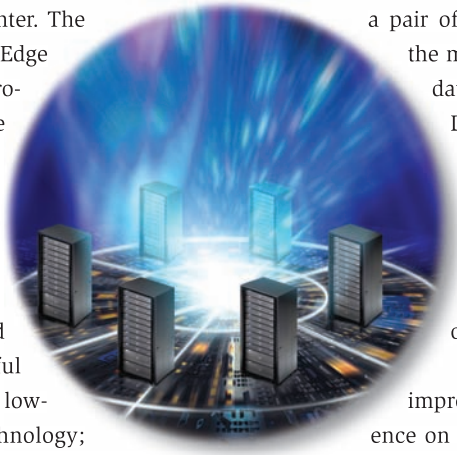
www.dell.co.uk/didyouknow

Easy as DELL



Power and Cooling 360

Recent advances in industry-standard server processor technology from AMD and Intel provide IT planners with attractive options when it comes to choosing technology for the energy-efficient data center. The Dell™ PowerEdge™ SC1435 and PowerEdge 6950 servers with AMD™ Opteron™ processors, as well as the Dell PowerEdge 1950 and PowerEdge 2950 servers with low-voltage Intel® Xeon® processors, offer a range of choices in the quest for improved performance per watt for key enterprise workloads.



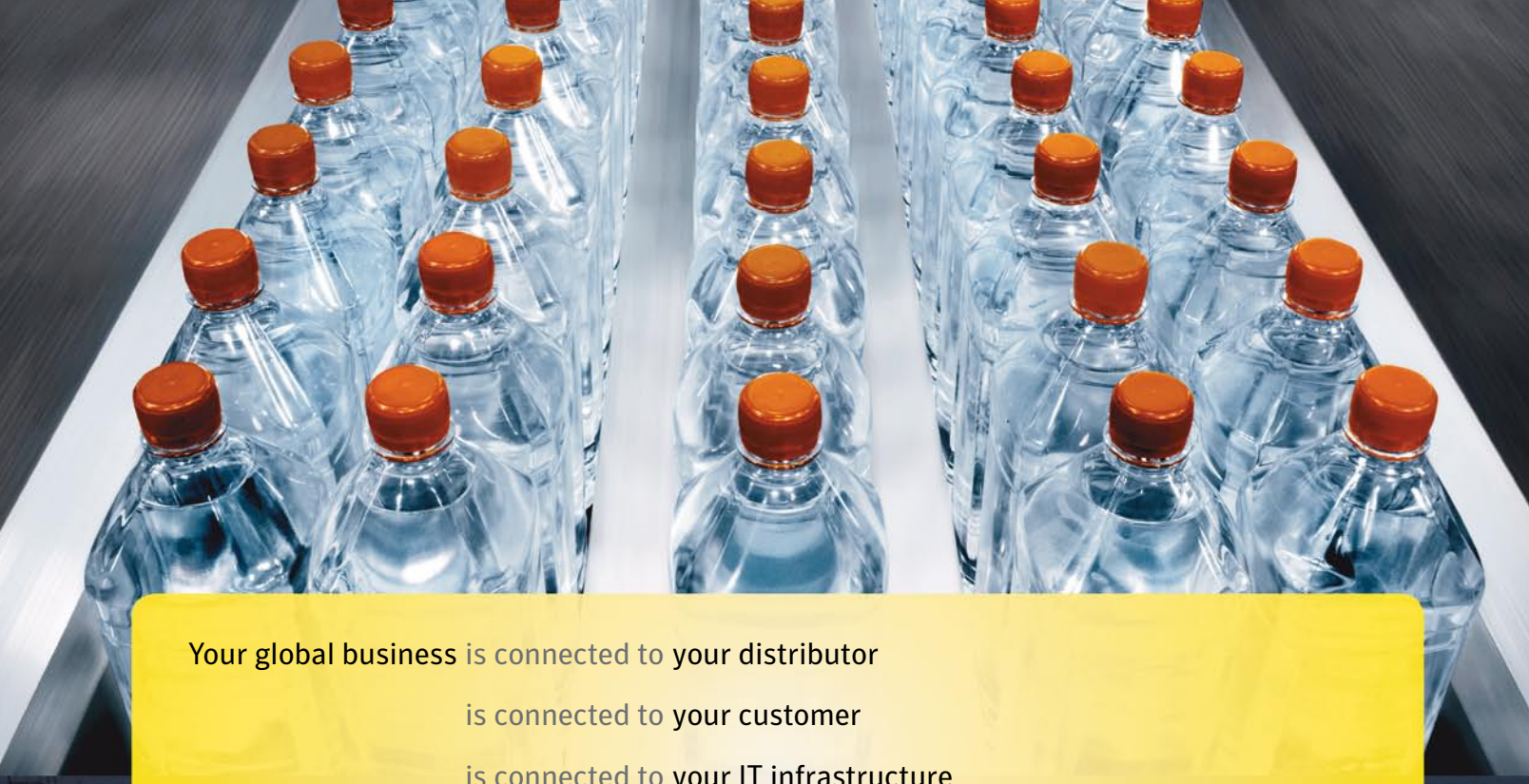
Power-smart features provided by these servers can include powerful multi-core processors that operate on low-voltage, innovative low-flow fan technology; high-efficiency power supplies; and special on-board firmware to help manage power consumption. And Dell's preconfigured Energy Smart servers can be an excellent way to gain access to these best-of-breed power features for your data center.

But there is much more to the power and cooling picture than purely performance per watt in the server farm. As we discovered when we began researching our special feature section on powering and cooling, IT planners would do well to consider a 360-degree perspective on data center efficiency. We again reached out to our Austin Design Center to gain insight into Dell's recent in-depth research on power consumption in a typical data center, and commissioned subject-matter experts from the Office of the Dell CTO—namely John Pflueger, Ph.D., and Sharon Hanson—to document the findings. The result was our cover story on page 8, "Data Center Efficiency in the Scalable Enterprise," which provides a holistic, yet pragmatic, perspective on managing power consumption across the entire data center at four levels: infrastructure, rack, individual server and storage system, and component.

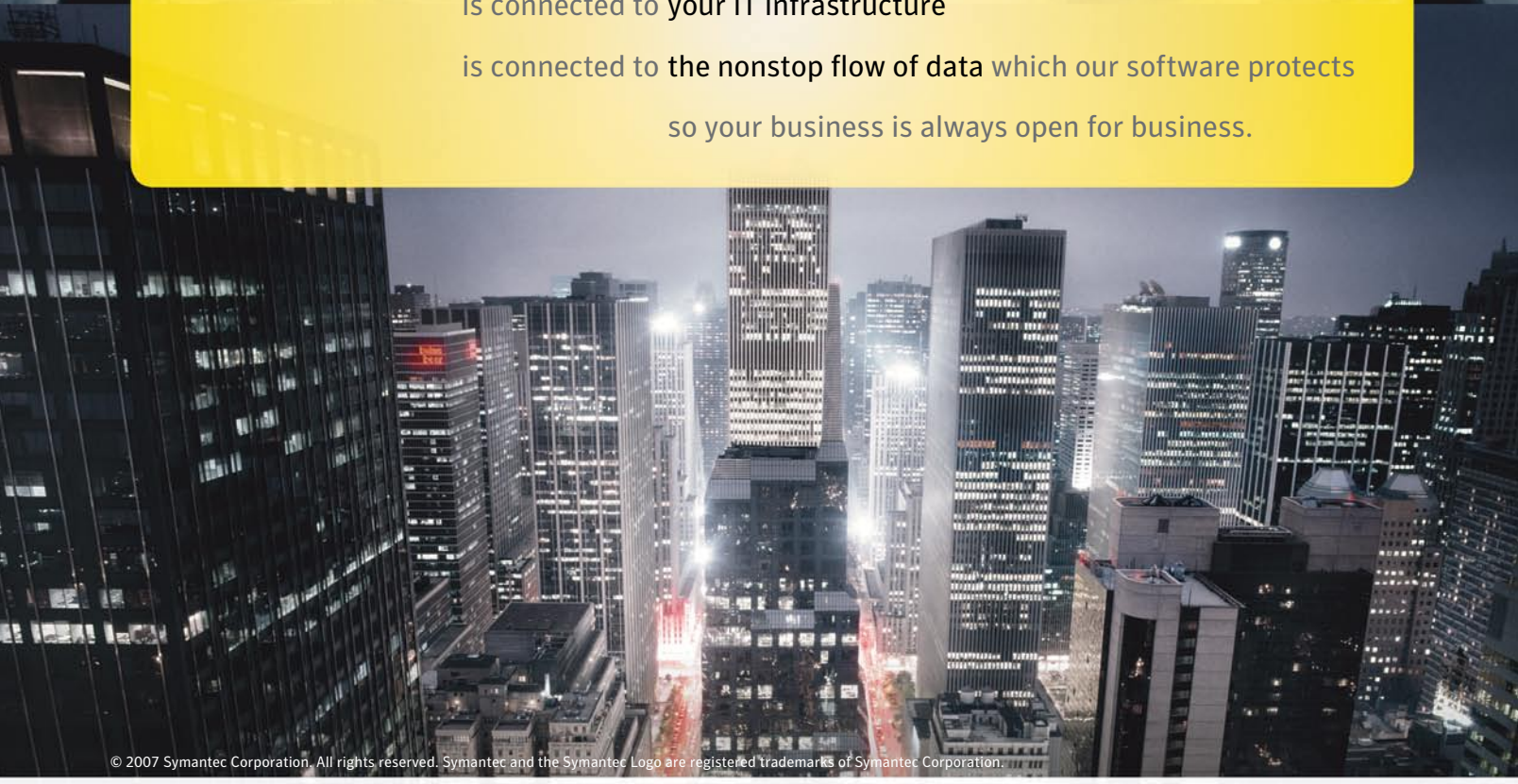
Additional coverage in the feature section includes a pair of articles that delve much deeper into the mechanics of power management in the data center for new AMD processor–based Dell servers: "Managing Data Center Power and Cooling with AMD Opteron Processors and AMD PowerNow! Technology" on page 16 and "Efficient Power Management on Dell PowerEdge Servers with AMD Opteron Processors" on page 22.

Meanwhile, we have been working to improve the efficiency of the user experience on the *Dell Power Solutions* Web site. For all newly posted article PDFs, you will see a "View Abstract" link just below the article's title, which activates an applet containing the title, byline, and full abstract. You can then elect to directly open the full PDF from the applet if the content is of interest to you. We hope you find this small enhancement helpful and the online content itself useful in your IT role.

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



Your global business is connected to your distributor
is connected to your customer
is connected to your IT infrastructure
is connected to the nonstop flow of data which our software protects
so your business is always open for business.

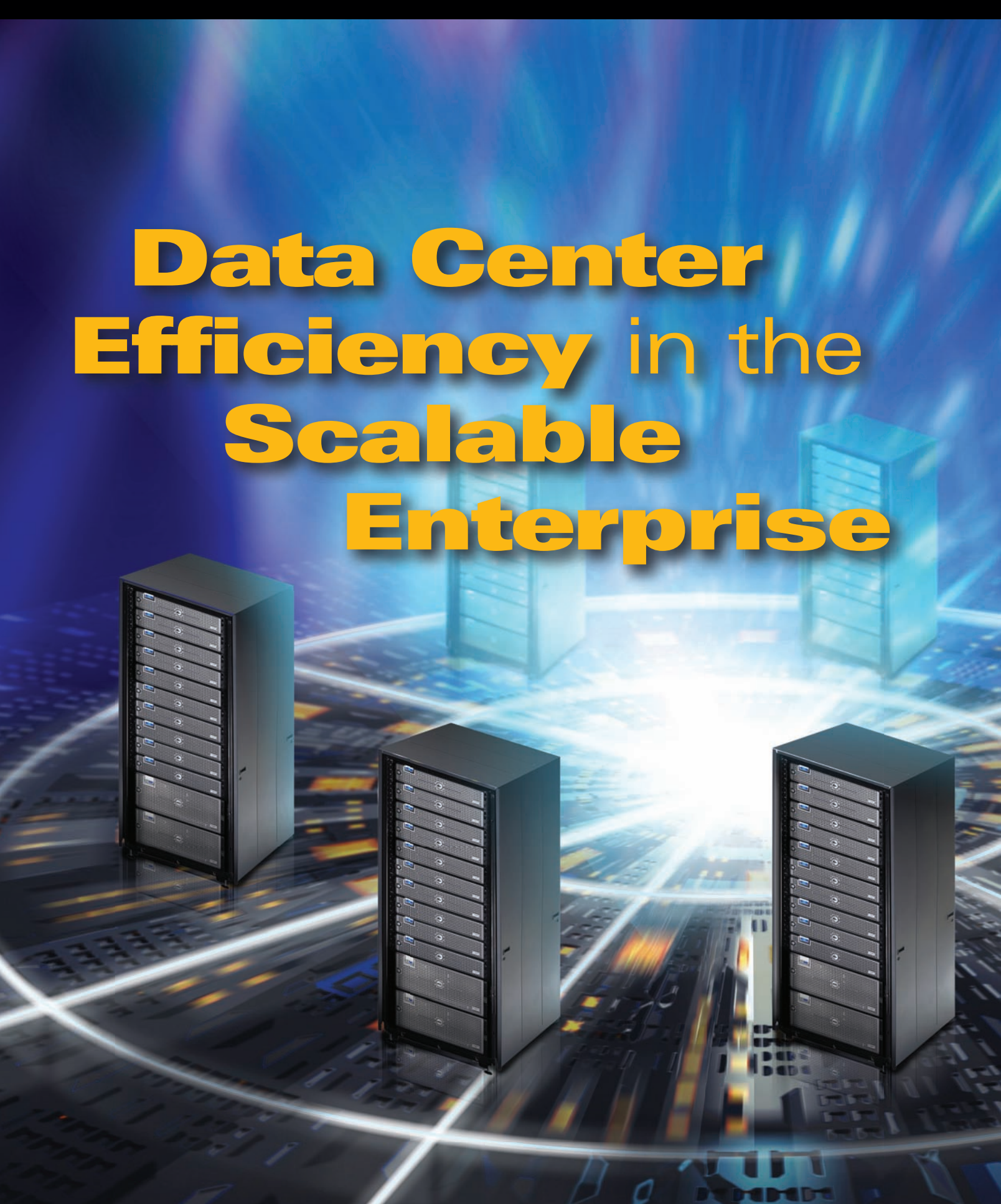


Software for the security, storage, and availability of your mission-critical information. These days, connections have never been so vital. Or so vulnerable. At Symantec, we have the solutions and expertise to keep your information secure and available across your entire operation. From backup and recovery to information security. From IT policy compliance to data center management. We bring confidence to every connection. Visit dell.com/symantec

Confidence in a connected world.

symantec™

Data Center Efficiency in the Scalable Enterprise



Related Categories:

Computational fluid dynamics (CFD)

Data center density

Dell PowerEdge servers

Power and cooling

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

Energy efficiency is a top concern as IT managers struggle with rising power bills, cooling problems, usage limits imposed by local utilities, or requirements to deploy additional servers without expanding an existing data center or building a new one. To address energy efficiency effectively, data center managers must look at a variety of issues ranging in scope from the smallest piece of silicon to the entire data center.

BY JOHN PFLUEGER, PH.D., AND SHARON HANSON

Looked at holistically, the data center life cycle offers two overarching opportunities for optimizing energy efficiency. First, enterprises can right-size the data center at the planning and deployment stage. Second, during the operating stage, enterprises can begin to take advantage of emerging trends toward more tightly managed power usage in the data center—from infrastructure to racks of servers to individual servers, and even components within those servers. Analyzing how power is used in a typical data center can be useful in understanding these opportunities.

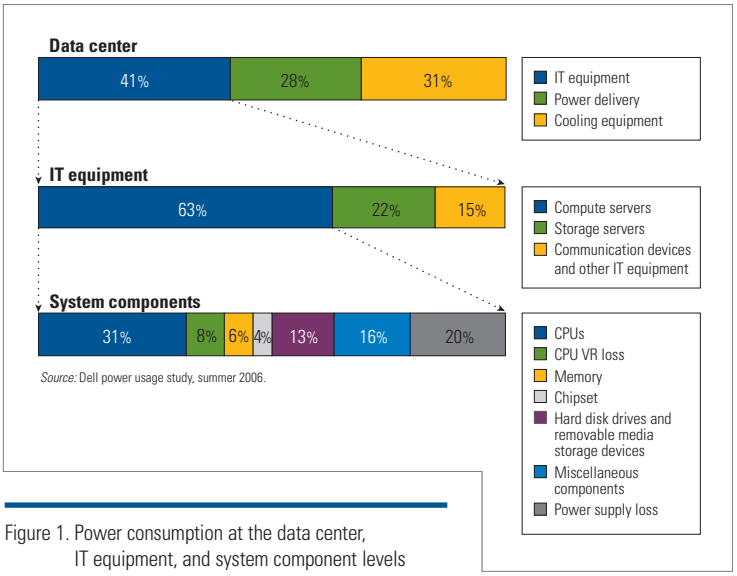


Figure 1. Power consumption at the data center, IT equipment, and system component levels

Measuring energy efficiency in the data center

A recent Dell study conducted during the summer of 2006 examined power consumption in a typical data center. This study looked at hour-by-hour power consumption at the data center level, then drilled down to the IT equipment and individual server levels. Figure 1 shows the results of the study.

At the data center level, the study showed that for every kilowatt delivered to the data center, 41 percent, or approximately 400 W, was delivered to the IT equipment. The remaining 59 percent was split between power distribution losses (28 percent) and the energy required to run cooling equipment (31 percent). This 41 percent efficiency is not atypical. In fact, a substantial number of data centers run at 30 percent efficiency or less.

Next, the study looked at how the 41 percent used by IT equipment was distributed among compute servers, storage servers, and other IT equipment. For this facility, roughly 63 percent of IT power was consumed by compute servers, 22 percent by storage servers, and 15 percent by the rest of the data center IT equipment.

The study then looked at the breakdown of average power usage of a common class of servers used in the data center. The results of this study show that the CPU used approximately 31 percent of the power, and approximately 28 percent of the power was consumed in power supply and CPU voltage regulator (VR) losses. The rest was split between memory, chipset, hard drives, removable media storage devices, and miscellaneous components.

This study shows that, although many IT organizations focus heavily on the power used by processors, addressing other infrastructure components provides at least as great an opportunity to save power and improve efficiency as focusing on processing power. For example, improving infrastructure efficiency from 41 percent up to 51 percent could allow the addition of 25 percent more servers to this data center.

This is a useful insight for those IT organizations challenged to increase performance per square foot in a space-constrained data center—improving infrastructure efficiency can be key. Such organizations may be constrained by infrastructure limitations or by utility providers that have cut them off at a certain power level. They are typically focused on increasing rack density so that they can produce more work in their existing space. However, the gating factor usually is not available rack space, but the ability to cool and provide power to the data center.

Building efficient data centers

To enhance data center efficiency, IT organizations must tightly manage power consumption across the entire data center at the infrastructure level, rack level, individual server and storage system level, and even component level within these systems. For power consumption in the data center, two best practices should be heeded:

- **Buy what you need:** In the planning and deployment stage of the data center life cycle, enterprises should right-size the data center.

METRICS FOR POWER EFFICIENCY

Today, power efficiency is measured as the ratio of power used by IT equipment to total data center power. Though useful in the past, when space and power typically were plentiful, the industry now needs more-meaningful ways to express efficiency that measure the amount of useful work produced. Efforts are underway in benchmarking organizations such as the Standard Performance Evaluation Corporation to develop such measures.

Today	Tomorrow
$\text{Power efficiency} = \frac{\text{Power}_{\text{IT equipment}}}{\text{Power}_{\text{Data center}}}$	$\text{Power efficiency} = \frac{\text{Useful work}_{\text{IT equipment}}}{\text{Power}_{\text{Data center}}}$

BEST PRACTICES: DATA CENTER EFFICIENCY

Data center planners and administrators should follow industry best practices to help improve data center efficiency. Following are some examples:

- Design for efficient cooling layout, such as implementing hot aisle/cold aisle configurations with suitably located air-conditioning equipment and installing blanking panels in unfilled server racks to preserve the hot aisle/cold aisle configuration.
- Right-size physical infrastructure systems to the anticipated workload.
- Put mechanisms in place to measure data center efficiency, and help ensure that both data center and facilities decision makers are accountable for data center efficiency.
- Refresh servers on a regular schedule to take advantage of the performance/watt benefits that accrue with succeeding generations.
- Prepare the organization for virtualization and begin planning for deployment of virtualization technology.

- **Use what you need:** In the operating stage of the data center life cycle, enterprises should take advantage of upcoming power management instrumentation to tightly manage power usages—from infrastructure to racks of servers to individual servers and components within those servers.

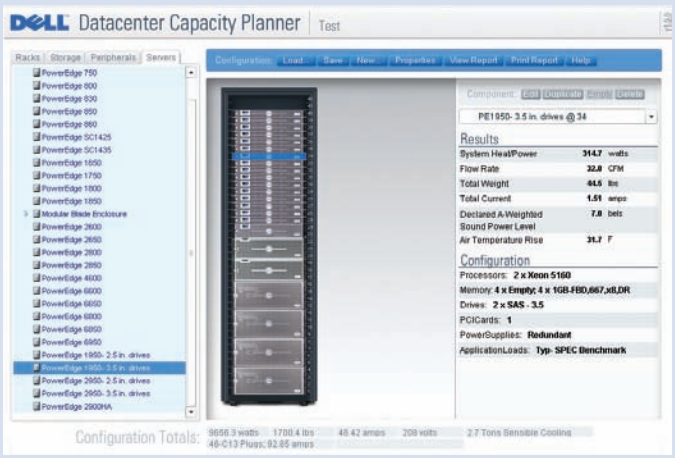
Buy what you need

The scalable enterprise is built upon a foundation of industry-standard data center building blocks, including storage systems and x86 servers that scale out in incremental, cost-effective phases to keep pace with business requirements for increased capacity and capability. The industry-standard x86 servers—equipped with one, two, or four CPU sockets—are interconnected so that the workload can be divided across multiple servers. Additional capacity is achieved by adding enough servers and associated storage to meet immediate growth requirements—thereby avoiding the big up-front investment in excess capacity that is often required when

DELL DATACENTER CAPACITY PLANNER

Dell™ products, tools, and services today help organizations throughout the data center life cycle: planning, deployment, operation, and refresh. Tools such as the Dell Datacenter Capacity Planner help estimate data center power and cooling requirements. Using this capacity planner, data center managers can configure racks of Dell servers, storage, and peripherals based on estimated power, cooling, and airflow for each rack and piece of equipment within the rack. To help improve data center efficiency, a suite of Dell IT infrastructure services include server and storage consolidation, data center environmental assessments, and virtualization readiness assessments. By partnering with industry leaders such as Liebert, Rittal, and Sanmina-SCI, Dell is working to address end-to-end data center efficiency with standards-based technologies.

For more information about the Dell Datacenter Capacity Planner, visit www.dell.com/energy.



Estimating power and cooling requirements using the Dell Datacenter Capacity Planner

scaling up with symmetric multiprocessor (SMP) servers that have eight or more CPU sockets.

Not only does the scale-out approach avoid the period of wasted capacity typical in the scale-up approach, it also avoids the extra power and cooling expense associated with running more compute capacity than is required. Excess capacity and overprovisioning are not just IT issues; they apply equally well to facility infrastructure. In fact, major infrastructure vendors have recently started introducing cooling and power products that embody this scale-out philosophy. For example, the Liebert Adaptive Architecture and the APC InfraStruXure architecture enable modular configurations for a variety of IT environments.

Use what you need

The pay-as-you-grow philosophy that underlies the scalable enterprise can be extended to using only what you need. For example, in a typical 24-hour period, servers may be operating at only 10–15 percent utilization more than 50 percent of the time. There is an opportunity for significant power savings if underutilized systems and even components within those systems can be transitioned to low-power states during nonpeak periods. A number of industry initiatives at the silicon, product, and infrastructure levels aim to reduce power consumption and to instrument data center IT equipment and infrastructure for more-granular power management than is currently available:

- **Silicon level:** A key focus of the IT industry is gaining power savings at the silicon level, particularly in processor, graphics, chipset, and memory components. This effort is beginning to show results. For example, major processor vendors AMD and Intel have both introduced power-saving features in their latest processors—Intel® Demand-Based Switching with Enhanced Intel SpeedStep® technology and AMD PowerNow!™ technology—and memory and chipset designs are beginning to reflect more power-efficient architectures.
- **Product level:** Dell product designs and specifications strike an aggressive balance between performance, features, and power consumption. In addition to supporting silicon vendors in their endeavors to create power-efficient architectures, Dell is requiring increasingly greater efficiencies in power supplies and voltage regulators, fans, and hard disk drives.

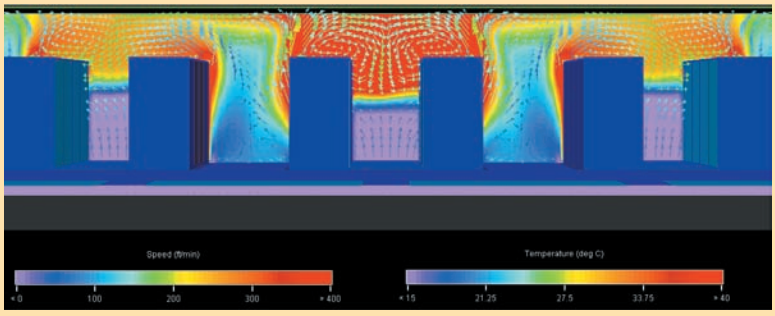
INSIDE THE DELL ENERGY EFFICIENCY RESEARCH CENTER

Thermal and cooling issues now rank among the top concerns at many enterprises. Over the past 10 years, server densities and power requirements in data centers have risen sharply: data centers once designed for approximately 30 W per square foot are now commonly designed for 120 W per square foot. This power is largely released as heat; in fact, servers release 100 percent of their input power as heat. As a result, the airflow required to cool increasingly dense server installations has begun to outstrip the capacity of typical raised-floor cooling systems. These systems generally handle 4–5 kW per rack, but a rack of blade servers today can require up to 25 kW.

To fully populate dense blade server racks, enterprises must consider cooling alternatives such as those installed in the Energy Efficiency Research Center at the Dell Austin Design Center. This 875-square-foot lab is designed to simulate an enterprise data center environment in which Dell engineers test and demonstrate innovative cooling techniques using current and emerging technologies such as liquid cooling at the chip level and liquid-cooled racks.

Heat and cooling simulation

The lab is densely populated with racks of special-purpose dummy servers—consisting of a large fan, heating core, processor, and control board—that enable Dell engineers to simulate different heating loads by setting specific fan, temperature, and power levels. The lab can produce more than 1,000 W per square foot of heat, which is well beyond today's average of 90–120 W per square foot.



The Dell team uses powerful 3-D computational fluid dynamics software to help address airflow and heat transfer issues. Shown here is a sample temperature and air vector plot for a small data center equipped with six rows of rack-mounted servers.

A smoke generator offers an effective mechanism for studying and demonstrating airflow: by showing recirculating air patterns, the lab team can test the effect of different cooling levels and techniques, and can visually demonstrate the thermal challenges organizations face as well as potential ways to mitigate those challenges. The lab is equipped with raised-floor cooling that is designed to provide up to three times the industry average in cooling capacity, overhead cooling that boosts capacity another 150 percent, and self-contained liquid-cooled racks.

Liquid-cooled racks

A liquid-to-liquid heat exchanger and pumping unit provides the interface between the building's chilled water and four self-contained cooling cabinets—the liquid-cooled racks. Each rack includes 20–30 kW of cooling capacity.

A liquid-cooled rack takes water from a closed secondary loop provided by the pumping unit. This water runs through coils in the left side of the front of the rack. The rack also generates its own airflow, so it does not rely on the raised floor for cooling: air is captured from the rear of the servers, pushed through these coils, and exhausted from the right side of the rack in front of the servers.

Ongoing thermal and cooling research

The Dell Energy Efficiency Research Center is dedicated to educating organizations and addressing cooling issues associated with increased server densities in enterprise data centers. It is also designed to scale to meet future requirements, including extra built-in cooling and electrical capacity that allow Dell engineers to model server loads significantly higher than those typically seen today. Such provisions equip the lab to continue testing and demonstrating innovative ways to optimize energy efficiency as power and cooling demands continue to rise.

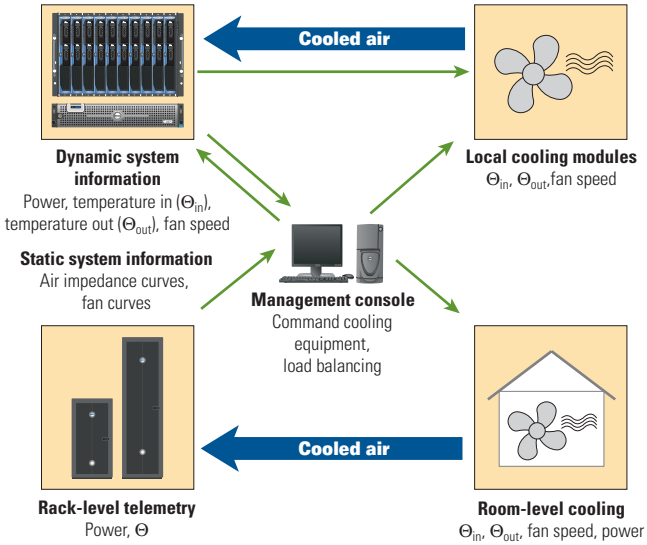


Figure 2. Power management in the data center

- **Infrastructure level:** Recent products from Liebert and APC also enable more-granular control over operations, providing cooling capacity that can be matched to the IT load. These products include modular uninterruptible power supply (UPS) systems and computer room air-conditioning equipment with variable-speed fans.

A clear trend is emerging: the data center ecosystem of IT, cooling, and power equipment is increasingly being instrumented for effective power management so that in the future the data center can be more tightly power-managed than is possible today. Eventually, virtualization technology can build on this capability to yield additional power management gains.

Enhancing data center efficiency through virtualization

Advances in virtualization technology have been fueled by multi-core processing, hardware support for virtualization in processors, multi-resource I/O initiatives, and low implementation costs. These advances are enabling mainstream adoption of virtualization technology. In the scalable enterprise data center, virtualization can be used to consolidate workloads from multiple underutilized servers to multiple virtual machines (VMs) on one server.

In addition, as systems management capabilities mature, virtualization can allow the data center workload—in the form of VMs—to be dynamically reallocated to meet changing business and demand requirements. This concept, sometimes referred to as workload mobility, can also be extended to power and thermal management. What if it were possible during nonpeak demand periods in the data center to consolidate the workload so that underutilized individual servers, racks of servers, or even entire segments of the data center could be transitioned to low-power states? And what if the associated power and cooling infrastructure were sufficiently manageable to also power down? Industry initiatives are underway that could lead to this level of tight power and thermal management in the data center.

Figure 2 shows what could become feasible when IT, power, and cooling equipment are instrumented for effective systems management. In this example, individual servers and racks provide static and dynamic system information such as power consumption levels, fan speed, airflow information, and temperature (represented by Θ) to a console. An increase in the load on a power supply could generate information for the console, allowing data center administrators to make intelligent decisions concerning rack or room-level cooling modules. In this way, administrators could manage the level of cooling for a server or server rack. More-comprehensive information about the state of the data center also could enhance data center managers' decision-making capabilities when planning and deploying new IT equipment. For example, IT equipment could be installed in locations known to have adequate cooling and power for the anticipated workload.

This vision, as exemplified in Figure 2, and the heightened interest in data center power consumption has led to collaborative industry efforts such as the Green Grid—an association of top IT companies, including Dell, that are committed to increasing energy efficiency in the data center.

Implementing data center efficiency in the scalable enterprise

Improving the efficiency of the data center requires looking at the decisions made during every phase of the data center life cycle. For example:

- **Engineering and planning:** Right-sizing the infrastructure for the expected IT load can provide an effective baseline for years of efficient IT operation.
- **Deployment:** Making smart choices about server location helps use available resources to best advantage.
- **Day-to-day operations:** Measuring and monitoring facility efficiency helps enforce discipline in the implementation of industry best practices.
- **Equipment refresh:** Replacing legacy systems with equipment that offers higher performance per watt than the systems being replaced enhances energy efficiency.

Dell's approach to energy efficiency addresses every aspect of the data center life cycle. Dell also remains focused on industry standards, leveraging interoperability standards where available and working to drive these standards where they are lacking. Problems such as energy consumption and density are only solved

ENERGY SMART DELL POWEREDGE SERVERS

The new Dell PowerEdge™ Energy Smart servers—the first industry-standard PowerEdge servers configured specifically to optimize energy efficiency—underscore Dell's commitment to driving energy-efficient solutions from the data center to the desktop. Available for PowerEdge 1950 and PowerEdge 2950 servers, Energy Smart technology helps reduce energy consumption and associated overall operating costs without sacrificing the performance required to run enterprise applications. Last fall, Dell announced the Energy Smart line of Dell OptiPlex™ desktops.

PowerEdge Energy Smart servers include several key features to help minimize power draw and cooling requirements:

- **Low-flow fan technology** automatically adjusts fan speeds based on environmental characteristics to avoid unnecessary power use.
- **High-efficiency power supplies** minimize wasted power and heat output to help reduce cooling needs.
- **Low-voltage Intel Xeon® 5148 processors** are designed to use reduced voltages and half the wattage of processors such as the Intel Xeon 5160.

By maximizing energy efficiency and minimizing cooling requirements, PowerEdge Energy Smart servers are designed to significantly increase performance/watt and reduce both data center operations costs and total cost of ownership. For more information, visit www.dell.com/energy/poweredge.

Energy Smart Dell PowerEdge 1950 (top) and PowerEdge 2950 servers help lower total energy consumption and reduce overall operating costs

through the development and adoption of standards from industry groups such as the Standard Performance Evaluation Corporation (SPEC); American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE); and Distributed Management Task Force (DMTF).

Going forward, scalability offers key opportunities to improve data center efficiency. First, administrators should avoid the inefficiencies involved in overprovisioning the data center. Second, administrators should leverage emerging technologies that can provide more-comprehensive power management capabilities. The focus on buying only what is required and using only what is needed in the data center is the fundamental thinking behind Dell's scalable enterprise strategy to simplify operations, improve resource utilization, and scale out cost-effectively.

John Pflueger, Ph.D., manages Dell's technical strategy for system and facility thermals, focusing on methods and system improvements for helping customers improve the efficiency of their computer systems and data centers. He has 15 years of experience in product development, product marketing, and product management. John has a Ph.D. from the Massachusetts Institute of Technology.

Sharon Hanson is a technical communications professional in the office of the Dell CTO. She has written and produced Dell white papers, articles, and other technical collateral on industry technology trends for the past 10 years. Sharon has a B.B.A. from the University of Texas at Austin.

FOR MORE INFORMATION

Dell energy efficiency:
www.dell.com/energy

BY PARTNERING WITH DELL, UNILEVER'S CUSTOMER SATISFACTION KEEPS GOING IN THE RIGHT DIRECTION. UP.

Unilever makes just about everything. Including making smart decisions, like choosing Dell™ Managed Services to get a higher return on their IT investment in North America. Unilever is now able to invest more in their brands, so their customers continue to feel good, look good and get more out of life. And Dell does it all by focusing on Unilever's agenda. That's how we design business solutions with only one company in mind. Yours.

To talk with a Dell representative call
1.866.212.9338

•

To learn more about the Unilever story visit
www.dell.com/unilever/ps

DELL Purely You

Managing Data Center Power and Cooling

with AMD Opteron Processors and AMD PowerNow! Technology

Avoiding unnecessary energy use in enterprise data centers can be critical for success. This article discusses the power and cooling advantages of AMD™ Opteron™ processors and AMD PowerNow!™ technology with Optimized Power Management, which are available in Dell™ PowerEdge™ servers.

BY BRENT KERBY

Related Categories:

AMD

AMD Opteron processors

Dell PowerEdge servers

Power and cooling

Power management

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

Today's data center is more than a room built to house rack upon rack of servers—it is a complex system. In fact, data centers have become so complex—with so many interactions among processors, rack systems, power and cooling systems, storage arrays, networks, and communications channels—that they can be regarded as ecosystems, requiring only the input of energy to become a world in themselves, while supporting virtually all critical business and scientific computing needs both inside and outside an enterprise.

The escalating speed, capacity, and power density of data center components have increased the degree of interdependence typically required for maximum performance and cost-effectiveness. Processor power, chip density, and the resulting power consumption and heat can have a profound effect on data center operational costs. And with data center managers under pressure to contain budgets and total cost of ownership, managing these effects can be critical to enterprise success.

AMD has developed several technologies to address this issue and has capitalized on its traditional strengths in low-power, high-performance processors; integration; and power management. The result: AMD Opteron processors and AMD PowerNow! technology with Optimized Power Management (OPM), which combine high-efficiency processing with sophisticated granular power management at the processor level. This article examines the underlying technologies of these products and how combining them can enhance efficiency and performance to help reduce data center costs.

Power and cooling challenges

Power consumption for cooling has long been regarded as a fixed cost. However, intelligent cooling system design,

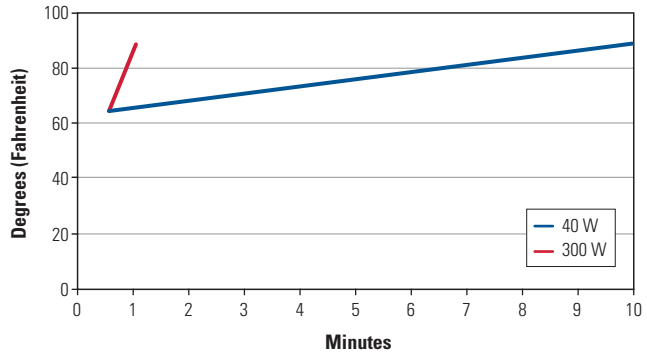


Figure 1. Heat density and thermal rise time

coupled with systems that can adjust power consumption based on performance demands, can make cooling power a variable, manageable cost that avoids wasted energy. The raw costs of electrical power to run a data center facility can likewise be managed, because processors typically consume a significant percentage of total system power.

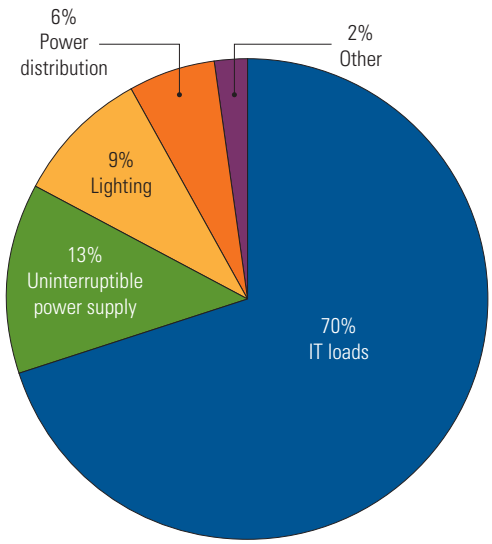
AMD takes these requirements as a responsibility to design and manufacture processors with the lowest possible power consumption and waste-heat generation. Maximizing existing data center facilities is critical to controlling total cost of ownership—data center managers cannot simply rebuild their infrastructure every time rack density or power consumption increases. Although typical racks installed in data centers just a few years ago might have consumed 2 kW of power and emitted 40 W of heat per square foot, in the coming years high-density racks may consume up to 25 kW per rack and dissipate as much as 500 W per square foot. This increased power density can cause much faster thermal rise than in older, lower-density systems: a data center with an average heat load of 40 W per square foot can cause a thermal rise of 25°F in 10 minutes, while an average heat load of 300 W per square foot can cause the same rise in less than a minute (see Figure 1).[1]

And because cooling is a major factor in data center construction costs, this type of increase can cause overall construction costs to skyrocket. For example, in 2004 it might have cost roughly US$400 per square foot to build a data center that emitted 40 W per square foot. Today, high-density racks producing up to 500 W per square foot can cost as much as US$5,000 per square foot for the necessary air conditioning, power supply, power conditioning, and other equipment. Thus a typical 50,000-square-foot facility that might have cost US$20 million a few years ago could cost as much as US$250 million today when factoring in projected future heat loads.[2]

Figure 2 shows where heat comes from in a typical data center. Because the majority of heat is generated by IT loads, reducing heat from these loads is a significant part of managing data center cooling in a cost-effective way: it can be significantly less expensive to replace blades and racks with reduced-heat, energy-efficient processors than to build new facilities to house unnecessarily hot components. To handle increased thermal rise times, data centers with high-density racks typically should include an uninterruptible cooling system; in addition, data-processing equipment must be self-regulating, highly responsive to thermal rise, and capable of reducing operating temperatures in a controlled way if the cooling system fails or becomes impaired. Given the heat density of modern rack systems, redundant, uninterruptible cooling systems are becoming increasingly popular. Changing a redundant cooling system from the prevalent $n + 1$ configuration to a $2n + 1$ configuration incurs costs, but it may be necessary to help reduce heat in the data center effectively.

AMD Opteron processors with AMD PowerNow! technology

AMD Opteron processors are designed to offer high performance and low power consumption. Their AMD PowerNow! technology with OPM uses dynamic frequency and voltage support, and has been refined since its introduction in June 2000 to deliver performance on demand while greatly reducing power consumption when full processor performance is unnecessary. In its latest, most advanced form, AMD PowerNow! technology with OPM under industry-standard Advanced Configuration and Power Interface (ACPI) program control allows processors to run at



Source: "Calculating Total Cooling Requirements for Data Centers," by American Power Conversion, www.apcmedia.com/salestools/NRAN-5TE6HE_R1_EN.pdf.

Figure 2. Relative contributions to total thermal output of a typical data center

[1] Bob Sullivan, senior consultant at the Uptime Institute, quoted in "Data Centers Get a Makeover," by Gary Anthes, in *Computerworld,* November 1, 2004, www.computerworld.com/printthis/2004/0,4814,97021,00.html.

[2] Ron Hughes, president of the California Data Center Design Group, quoted in "Data Centers Get a Makeover," by Gary Anthes, in *Computerworld,* November 1, 2004, www.computerworld.com/printthis/2004/0,4814,97021,00.html.

multiple frequencies and voltages without changing memory or frontside bus speeds.

Even without AMD PowerNow! technology, AMD Opteron processors can provide the type of efficient, low-heat processors for high-density server applications that help reduce overall cooling costs—the standard-power dual-core AMD Opteron processor, which includes an integrated memory controller, consumes only 95 W maximum. Adding AMD PowerNow! technology to this processor enhances energy-consumption performance by allowing multiple levels of reduced clock speed and voltage.

Optimizing processing power can provide numerous benefits for enterprise data centers; these benefits can increase exponentially as the number of processors rises, because of the increasing amount of electricity required to run the processors as well as the power to cool them and additional platform components. By enabling increased granularity and reduced overall power consumption, AMD technology allows enterprises to increase data center density without expensive retrofitting of high-capacity cooling systems.

Efficient design

Dual-core AMD Opteron processors can enhance performance, efficiency, and cost-effectiveness in enterprise data centers. Having multiple identical processor cores on the same chip offers several important advantages:

- **Low power consumption:** The "edge" components of a processor—drivers, voltage regulators, bus interfaces, and the like—use a substantial amount of its power. Sharing these components between two cores uses less power than if they were duplicated in two separate processors.
- **High interconnect speeds:** Chip-level interconnects are inherently faster than chip-to-chip interfaces, yielding a significant performance advantage. The on-chip system request queue that mediates traffic to and from the cores is an integral part of the overall design.
- **High throughput with reduced power consumption:** AMD PowerNow! technology P-states are automatically synchronized between cores, allowing the workload to be shared transparently and providing increased throughput with reduced power consumption compared to separate processors.
- **High performance per watt:** Dual-core AMD Opteron processors use a common socket that is planned to support future quad-core AMD Opteron processors and be able to run in many systems, requiring nothing more than a BIOS upgrade. Quad-core AMD Opteron processors are intended to fall within the same thermal envelope as dual-core AMD Opteron processors while providing increased performance, resulting in increased performance per watt over dual-core AMD Opteron processors.

The Direct Connect Architecture used in AMD Opteron processors connects the processing cores directly to the integrated memory controller, I/O, and other processors, which helps reduce chip count, system bottlenecks, and power consumption. The integrated memory controller effectively eliminates the traditional Northbridge controller and thus provides several benefits, the most important of which is enhanced performance because the processor no longer has to communicate over a frontside bus to its memory controller or I/O. Because the memory controller is an integral part of the chip, there is no bus contention for the controller—the crossbar switching built into the AMD Opteron processor gives both cores easy access to the memory controller.

AMD PowerNow! technology provides sophisticated processor power management by establishing multiple combinations of processor voltage and frequency called P-states.

Every piece of silicon on a server circuit board costs money and adds heat. Each additional device that requires a board- or bus-level interconnect is a bottleneck—not a potential bottleneck, but one with a measurable effect on overall system throughput. The savings in board space, power consumption, and overall system cost are clear and quantifiable. By integrating the memory controller into the same chip that holds the two cores in the dual-core AMD Opteron processor, AMD has eliminated bus interface chips, synchronizing signals, and similar overhead, saving tens of watts per system along with the related reduced heat load. These advantages can add up quickly, particularly in rack systems.

And because AMD Opteron processors were designed from their inception to be multi-core capable, a future transition from dual-core to quad-core processors is planned to be simple and transparent, requiring nothing more than a BIOS upgrade. AMD plans for future quad-core AMD Opteron processors to operate within the same power and thermal envelopes as the dual-core models while providing increased performance, helping deliver optimal performance per watt.

Sophisticated, OS-independent power management

AMD PowerNow! technology provides sophisticated processor power management by establishing multiple combinations of processor voltage and frequency called P-states. Although these states can be defined and managed in the BIOS, multiprocessor systems implement them in the OS kernel. AMD Opteron processors can respond very quickly—in microseconds—to P-state changes, enabling processor power and speed to follow workload closely. The highest P-state runs the processor at full clock speed and full voltage, but during off-peak

It's time for you to lay off the juice.

The energy-efficient AMD Opteron™ processor.

AMD
Smarter Choice

www.amd.com/lessenergy

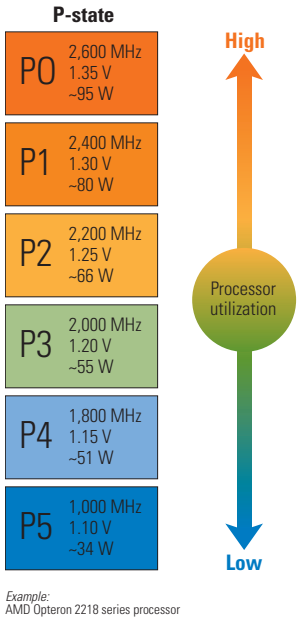


Figure 3. Power and cooling management with processor performance states

conditions, the clock can drop all the way back to a 1 GHz "idle," saving as much as 75 percent of the full-speed power, and make continuous granular changes between these two states (see Figure 3). This technology helps avoid unnecessary heat buildup, even in extremely discontinuous workload scenarios.

AMD PowerNow! technology is implemented using the ACPI standard, which was originally developed for notebook power management but has become a general power management scheme with inherent multiprocessor support and platform independence. Using industry-standard ACPI calls allows AMD PowerNow! technology to interoperate with cooling systems, power systems, and facilities management software and with many different operating systems, including Microsoft® Windows® and Linux® operating systems.

Microsoft Windows XP. Windows XP has several power-down states, including global-level states, sleep states, device-level states, and four dedicated C-states for the processor. Within its highest-power (or *normal*) C-state, C0, the processor driver can impose several subdivided *performance states.* AMD PowerNow! technology is implemented in this driver, and can operate independently of Windows XP as long as the OS has invoked the C0 state. Although it is somewhat unusual for server operating systems to invoke sleep states, if the OS were to go into standby, it would instruct the processor to go into its C3 state, which uses the lowest amount of power of the C-states.

Windows XP organizes its power management directives into *policy* and *non-policy* groups, which communicate with the AMD Opteron processors through ACPI. These directives map well to AMD PowerNow! technology capabilities. The *Adaptive* policy, for example, reduces processor performance to the lowest voltage and frequency state available whenever processor demand does not justify a higher state. This policy does not utilize linear stop-clock throttle states, except in response to thermal events.

Non-policy states, in contrast, are exceptions that, within the scope of server applications, are designed to conserve power in emergencies or prevent thermal damage. If the temperature exceeds a passive thermal point stored as a registry value, the OS uses successively lower performance levels to reduce the temperature below this point. If reduced performance levels do not alleviate the problem, the kernel uses stop-clock throttling to help prevent damage.

Linux. Linux has proactive OS-directed power management (OSPM) cooling policies and a comprehensive ACPI implementation, with table-driven settings that map to AMD PowerNow! technology P-states. Although OSPM can poll thermal zones, typical implementations use asynchronous signaling to eliminate wasteful polling overhead. Like Windows power management, OSPM encompasses both active (for performance) and passive (for energy conservation) cooling models and implements critical points with multiple thermal thresholds for system protection and orderly idling or shutdown in case of thermal failure.

Potential benefits of AMD Opteron processors and AMD PowerNow! technology

AMD Opteron processors and AMD PowerNow! technology with OPM can provide the following benefits:

- Optimized computational power per watt
- Reduced heat dissipation per unit of computational power
- Enhanced granular power management to help minimize power consumption and excess heat
- Industry-standard power management interfaces for easy integration into total data center management
- Significant performance increases over equivalent single-core processors

Data center managers should not overlook this last benefit: a dual-core processor typically can provide nearly twice the processing power of a single-core processor without increasing the heat load (and possibly even reducing it). And because some software vendors—Microsoft, for example—maintain the same license prices for dual-core processors as for single-core processors, enterprises can gain increased processing power without attendant increases in licensing costs.

AMD Opteron processors and AMD PowerNow! technology with OPM are designed in particular to provide two benefits: efficient power use and enhanced performance.

Efficient power use

Data center servers typically have fluctuating workloads throughout the day, causing variable processor utilizations depending on demand. Understanding the average utilization of their installed server base and the power consumption across multiple utilization levels can give data center managers an indication of possible power savings in a particular environment.

Figure 4 illustrates how AMD PowerNow! technology can minimize processor power consumption as load decreases, and how OPM can maximize the potential benefits: AMD PowerNow! technology

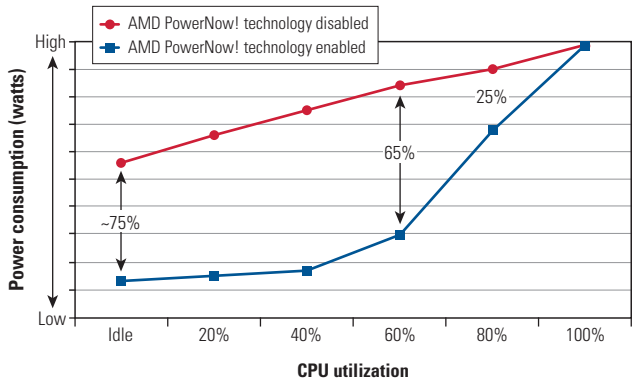


Figure 4. Average CPU core power measured with AMD Synthetic CPU Workload Utility

can provide platform-level power savings ranging from 25 percent at 80 percent processor utilization to 65 percent at 60 percent processor utilization to approximately 75 percent during processor idle.[3]

AMD PowerNow! technology with OPM matches low-load times with low-level P-states to help avoid unnecessary power use. Virtualization, for example, is vitally important in many data centers—and in these environments, virtualized processors, storage volumes, drive spindles, processes, and communication channels are constantly changing. Applications may demand additional resources as workloads increase, but ACPI, which functions at the processor level, can manage power consumption even as conditions change dynamically. Deploying AMD PowerNow! technology with OPM in conjunction with hot aisle/cold aisle layouts[4] and retrofitting with new racks and blades can provide multiple benefits, including allowing data centers to preserve investments in existing power and cooling infrastructures, reducing overall cooling power consumption and power consumption per server, and increasing processor density.

Enhanced performance

One of the truisms in IT is that there is no such thing as too much bandwidth—and in the server world, today's acceptable bandwidth is tomorrow's bottleneck. AMD has made specific enhancements to the AMD Opteron cache coherency algorithms that are designed to provide each processor extremely rapid and efficient access to other processors' cache contents, even though each processor maintains its own 1 MB level 2 (L2) cache. Because the majority of data center applications are inherently multi-threaded, they can benefit greatly from this architecture. Memory bandwidth can be more important than core clock speed when properly applied, with processors coupled closely through crossbar switching and a memory controller that runs at the processor's speed, as in AMD Opteron processors.

AMD Opteron processors can also provide performance advantages for communicating with memory and other I/O. Because the memory controller is integrated with the chip and runs at the processor's speed, communications between the memory controller and processors avoid frontside bus bottlenecks, which can result in significantly lower latencies than a traditional external memory controller. Multiple AMD Opteron processors communicate with one another through AMD HyperTransport™ links running at 1 GHz, with a total theoretical bandwidth for AMD Opteron 2000 and 8000 series processors of 32 GB/sec. These speeds are critical for processor-to-processor communication, particularly for the AMD Opteron 8000 series processor, which is optimized for four-way and eight-way configurations.

With so many applications executing 32-bit code, it can be tempting to dismiss the 64-bit capabilities of AMD Opteron processors as a "someday" feature. But 64-bit operating systems, compilers, and applications are increasingly available, and depending on the workload, can be dramatically faster than their 32-bit counterparts for compute-intensive tasks. Data center managers should keep in mind that the difference between 32 and 64 bits is not only the number of bits—64-bit systems provide exponentially more memory that can be directly addressed than 32-bit systems.

Efficient power and cooling in enterprise data centers

For data center managers, the increasing demands, processing requirements, rack density, heat loads, energy consumption, and costs for new and upgraded cooling equipment create greater pressure than ever before to make wise technology investments. Dual-core AMD Opteron processors with AMD PowerNow! technology and Optimized Power Management provide a low-power, high-efficiency option that helps reduce power consumption and cooling costs in enterprise data centers.

Brent Kerby is the product marketing manager for AMD Opteron processors. He has more than 20 years of experience in both technical and marketing positions for consumer and commercial computing products. Brent is currently focused on elements of data center total cost of ownership, including performance-per-watt efficiency and power management.

FOR MORE INFORMATION

AMD Opteron processors:
www.amd.com/opteron

Dell systems featuring AMD processors:
www.dell.com/amd

[3] These average results are based on a four-socket internal AMD test platform with four AMD Opteron 8220 SE processors; four 1 GB, 667 MHz double data rate 2 (DDR2) dual in-line memory modules (DIMMs) per socket (for a total of 16 GB of memory); and a 250 GB Serial ATA (SATA) hard drive.

[4] Hot aisle/cold aisle layouts help maximize the efficiency and effectiveness of cooling systems, and lend themselves well to spot-cooling with auxiliary fans or small refrigeration units to direct additional cold air to a particular piece of equipment. For more information, see "Alternating Cold and Hot Aisles Provides More Reliable Cooling for Server Farms," by the Uptime Institute, www.upsite.com/TUIpages/whitepapers/tuiaisles.html.

Efficient Power Management

on Dell PowerEdge Servers with AMD Opteron Processors

Efficient power management enables enterprises to help reduce overall IT costs by avoiding unnecessary energy use. This article describes how to enable and validate AMD PowerNow!™ power management technology on Dell™ PowerEdge™ servers with AMD™ Opteron™ processors.

BY MANOJ GUJARATHI, LISA GARRISON, AND JORDAN HARGRAVE

Related Categories:

AMD

AMD Opteron processors

Dell PowerEdge servers

Microsoft Windows Server 2003

Novell SUSE Linux

Power and cooling

Power management

Red Hat Enterprise Linux

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

Efficient power management in enterprise data centers has become increasingly critical to help control costs. Dell PowerEdge servers with AMD Opteron processors enable enterprises to take advantage of AMD PowerNow! technology to provide this management in their data centers. AMD PowerNow! technology allows operating systems to dynamically adjust processor power states, voltage, and clocking frequencies depending on workload, which helps reduce total power consumption and related IT costs while providing enhanced performance/watt capabilities and performance on demand; for example, the processor can run at a low power state when the system is idle, avoiding unnecessary power use and reducing cooling requirements. AMD tests have shown that enabling AMD PowerNow! technology allows for power savings of up to 65 percent at 60 percent processor utilization and up to 75 percent during processor idle when compared to power consumption with this technology disabled.[1]

This article provides step-by-step guidance to help administrators enable AMD PowerNow! technology on Dell PowerEdge servers with AMD Opteron processors, including details on additional software and driver components required to use AMD PowerNow! technology as well as basic ways of validating its functionality to help ensure effective power management. The instructions in this article apply to PowerEdge 6950 and PowerEdge SC1435 servers (see the "Power-friendly Dell PowerEdge servers with AMD processors" sidebar in this article) with AMD Opteron processors running the Microsoft® Windows Server® 2003 OS with Service Pack 1 (SP1), Microsoft Windows Server 2003 x64 Editions, the Red Hat® Enterprise Linux® OS, or the Novell® SUSE® Linux Enterprise Server OS.

[1] These average results are based on a four-socket internal AMD test platform with four AMD Opteron 8220 SE processors; four 1 GB, 667 MHz double data rate 2 (DDR2) dual in-line memory modules (DIMMs) per socket (for a total of 16 GB of memory); and a 250 GB Serial ATA (SATA) hard drive. For more information about AMD PowerNow! technology and processor utilization, see "Managing Data Center Power and Cooling with AMD Opteron Processors and AMD PowerNow! Technology," by Brent Kerby, in *Dell Power Solutions*, February 2007, www.dell.com/downloads/global/power/ps1q07-20070204-AMD.pdf.

When implementing a power management system, enterprises should always consider the trade-off between power efficiency and performance, particularly the impact that power management can have on the performance of critical applications.

Updating the BIOS to enable power management

The first step in using AMD PowerNow! technology on Dell servers is to update the BIOS token to enable power management. Administrators can do this by changing the Demand-based Power Management BIOS token from "Disabled" to "Enabled."

In addition to this manual option, administrators can use the Dell OpenManage™ Deployment Toolkit (DTK), Dell OpenManage Server Administrator (OMSA), or the BIOS setup utility to update the token. These tools can play a key role when carrying out one-to-one (local or remote) or one-to-many BIOS token updates. For all of these methods, a system reboot is required to make the token update effective.

Dell OpenManage Deployment Toolkit. Administrators can use the following command in the DTK to update the Demand-based Power Management BIOS token:

```
syscfg --dbpm=enable
```

This command can be replicated for one-to-many environments or scripted.

Dell OpenManage Server Administrator. Administrators can use the OMSA command-line interface (CLI) or graphical user interface (GUI) to carry out one-to-one system monitoring and configuration. To update the Demand-based Power Management BIOS token through the CLI, administrators can use the following command:

```
omconfig chassis biossetup
   attribute=dbpm setting=enable
```

To update the token through the GUI, administrators can select the appropriate server, go to System > Main System Chassis > BIOS, select the Setup tab, then select "Enable" (see Figure 1).

BIOS setup utility. Administrators can update the Demand-based Power Management BIOS token with the BIOS setup utility by performing the following steps:

1. Power the system on.
2. Press F2 during the system power-on self-test to enter the BIOS setup utility.
3. Press the right-arrow key to reach the Advanced screen.
4. Press the down-arrow key to highlight "Power Management" and press Enter.
5. Press the right-arrow key to change the mode to "Enable."
6. Exit the utility and allow the system to continue booting.

Enabling AMD PowerNow! technology in Microsoft Windows Server 2003

After updating the BIOS, administrators must perform three key steps to enable AMD PowerNow! technology on servers running Windows Server 2003 with SP1 or Windows Server 2003 x64 Editions: installing the AMD PowerNow! driver, installing the appropriate OS hot fix, and selecting the appropriate power scheme.

Installing the AMD PowerNow! driver

Using AMD PowerNow! technology with Windows Server 2003 requires version 1.3.2 or later of the amdk8.sys driver. Administrators can download the latest version of this driver from www.amd.com.

After installing the driver, administrators must modify the boot.ini file to include the `/USEPMTIMER` switch for multiprocessor support. This switch forces the system to use the power management timer (PMT) rather than the time stamp counter (TSC). Because of the frequency-variant nature of TSCs, using TSCs instead of PMTs can cause symmetric multiprocessing (SMP) system processors to run at different frequencies.

Administrators can avoid this step by using the AMD setup utility to install the driver; the installer automatically updates the boot.ini file to include the `/USEPMTIMER` switch. They can verify this switch is present by checking the boot.ini file following driver installation.

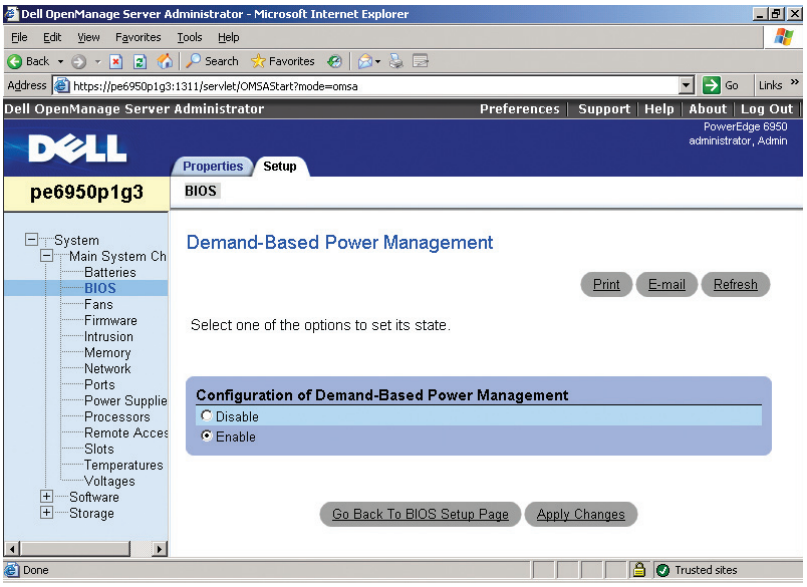


Figure 1. Demand-based Power Management screen in the Dell OpenManage Server Administrator GUI

POWER-FRIENDLY DELL POWEREDGE SERVERS WITH AMD PROCESSORS

Dell PowerEdge 6950 and PowerEdge SC1435 servers are designed for efficiency and scalability. Dual-core AMD Opteron processors enable these servers to take advantage of both the efficiency of the AMD Direct Connect Architecture and the power management capabilities of AMD PowerNow! technology, helping reduce overall data center power and cooling costs.

The PowerEdge 6950 server uses dual-core AMD Opteron 8200 series processors. These 64-bit processors help PowerEdge 6950 servers provide high performance while minimizing power consumption. For example, tests performed by Dell engineers in August 2006 showed that the PowerEdge 6950 used up to 20 percent less power than a previous-generation PowerEdge 6850 server configured with dual-core Intel® Xeon® 7100 series processors.*

The PowerEdge SC1435, designed for high-performance computing clusters and distributed Web serving, uses dual-core AMD Opteron 2200 series 64-bit processors that help provide exceptional performance and performance per watt compared with previous-generation PowerEdge SC servers. For example, tests performed by Dell engineers in September 2006 demonstrated that the PowerEdge SC1435 could provide up to 128 percent higher performance and up to 138 percent higher performance per watt than a previous-generation PowerEdge SC1425 server with single-core Intel Xeon processors.**

Dell PowerEdge 6950 servers enable exceptional performance for demanding applications such as database, server consolidation, virtualization, and migration from costly proprietary RISC-based systems

Because they are based on industry-standard components, Dell PowerEdge servers enable organizations to simplify operations, improve resource utilization, and scale out in cost-effective increments. In addition, the dual-core AMD processor–based servers are designed to provide electrical, thermal, and socket compatibility with upcoming quad-core AMD processors. Once those processors become available, enterprises should be able to seamlessly upgrade dual-core AMD processor–based systems for increased performance.

For more information about Dell systems with AMD processors, visit www.dell.com/amd or www.dell.com/energy.

Dell PowerEdge SC1435 servers provide an excellent platform for high-performance computing clusters, file and print sharing, and Web farms

* Based on AC power measurements using an Extech 380803 Power Analyzer taken during the peak load of the SPECjbb2005 benchmark test performed by Dell Labs in August 2006 on a PowerEdge 6950 with four dual-core AMD Opteron 8220 SE processors at 2.8 GHz, 32 GB of 667 MHz DDR2 memory, and Microsoft Windows Server 2003 Enterprise x64 Edition with SP1, as compared with a PowerEdge 6850 with four dual-core Intel Xeon 7140 processors at 3.4 GHz, 16 GB of 400 MHz DDR2 memory, and Windows Server 2003 Enterprise x64 Edition with SP1. Actual power consumption will vary based on configuration, usage, and manufacturing variability.

** Based on SPECfp_rate benchmark tests performed by Dell Labs in September 2006. Performance tests used a PowerEdge SC1435 server with two AMD Opteron 2220 SE processors at 2.8 GHz; 8 GB of 667 MHz DDR2 memory; a SATA 80 GB, 7,200 rpm hard drive; and the 64-bit Red Hat Enterprise Linux 4 Update 4 OS. Performance/watt tests used a PowerEdge SC1435 server with two AMD Opteron 2218 processors at 2.6 GHz; 8 GB of 667 MHz DDR2 memory; a SATA 80 GB, 7,200 rpm hard drive; and the 64-bit Red Hat Enterprise Linux 4 Update 4 OS. Both were compared with a PowerEdge SC1425 with two single-core Intel Xeon processors at 3.8 GHz; 8 GB of 400 MHz DDR2 memory; a SATA 80 GB, 7,200 rpm hard drive; and the 64-bit Red Hat Enterprise Linux 4 Update 4 OS. Actual performance and power consumption will vary based on configuration, usage, and manufacturing variability.

Installing the appropriate OS hot fix

To avoid a known deadlock that can occur in the Windows Server 2003 kernel on systems with two or more AMD Opteron processors, administrators should install the hot fix available at support.microsoft.com/kb/924441.

Selecting the appropriate power scheme

Administrators should set the server to use the "Server Balanced Processor Power and Performance" power scheme. They can set this option by selecting Start > Control Panel > Power Options and using the "Power schemes" drop-down menu, as shown in Figure 2.

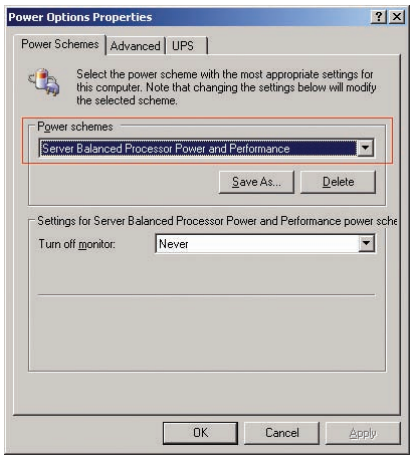


Figure 2. Power scheme selection in Microsoft Windows Server 2003

Enabling AMD PowerNow! technology in Linux operating systems

Support for AMD PowerNow! technology in Red Hat Enterprise Linux 4 Update 4 or later and Novell SUSE Linux Enterprise Server 10 is provided by the powernow-k8 driver. Figure 3 summarizes Linux support for AMD PowerNow! technology on PowerEdge servers with AMD Opteron processors. Red Hat Enterprise Linux 4 prior to Update 4 does support dynamic processor-frequency drivers; however, the version of the AMD PowerNow! driver included with the OS does not detect the AMD Opteron processors used in PowerEdge servers, so administrators should disable AMD PowerNow! technology in the BIOS when using Red Hat Enterprise Linux 4 prior to Update 4.

To enable AMD PowerNow! technology on systems running these Linux operating systems, administrators must first check that the system correctly detects the AMD PowerNow! driver, and load it if necessary. In Red Hat Enterprise Linux, they must also carry out some additional configuration steps. Administrators should always check that they are using the latest driver versions, which are available for download from www.amd.com.

Detecting the AMD PowerNow! driver

Administrators can determine whether a Linux-based system is correctly detecting the AMD PowerNow! driver with the following command:

```
dmesg | grep -i powernow
```

If this command does not produce any output, they should download the latest version of the powernow-k8 driver from www.amd.com, then load it with the following command:

```
modprobe powernow-k8
```

Figure 4 shows the two types of output this command produces depending on whether demand-based power management is enabled in the BIOS. If the command produces output indicating that this option is not enabled, administrators should follow the steps in the "Updating the BIOS to enable power management" section in this article. If AMD PowerNow! technology is enabled in the BIOS, the command loads the driver, then displays the speeds and voltages the processor supports.

Enabling AMD PowerNow! technology in Red Hat Enterprise Linux

To enable AMD PowerNow! technology in Red Hat Enterprise Linux 4, administrators must also perform the following steps:

1. In /etc/cpuspeed.conf, edit or add the `DRIVER` line to read `DRIVER="powernow-k8"`.
2. Enable the cpuspeed daemon with the command `chkconfig cpuspeed on`.
3. Load the powernow-k8 driver with the command `modprobe powernow-k8`.
4. Start the cpuspeed daemon with the command `/etc/init.d/cpuspeed start`.

Using processor frequency tools

Several tools are available in Red Hat Enterprise Linux and SUSE Linux Enterprise Server to adjust processor frequency settings. Red Hat Enterprise Linux primarily uses the cpuspeed daemon, and SUSE Linux Enterprise Server primarily uses the powersave utility.

	Powernow-k8 driver version	i386 support	x86-64 support
Red Hat Enterprise Linux 4 prior to Update 4	1.00.09b	No (module)	No (built-in)
Red Hat Enterprise Linux 4 Update 4	1.60.0	Yes (module)	Yes (built-in)*
Novell SUSE Linux Enterprise Server 10	1.60.2	N/A	Yes

* For the x86-64 version of Red Hat Enterprise Linux 4 Update 4, the powernow-k8 driver is included in the kernel and cannot be updated without a kernel recompile.

Figure 3. Support for AMD PowerNow! technology in Red Hat Enterprise Linux and Novell SUSE Linux Enterprise Server

In Red Hat Enterprise Linux, administrators can use the following commands to set the processor frequency scaling method:

`killall -SIGUSR1` (processor runs at its fastest speed)
`killall -SIGUSR2` (processor runs at its slowest speed)
`killall -SIGHUP` (processor speed is adjusted dynamically)

In SUSE Linux Enterprise Server 10, the powersave utility allows administrators to change several power-related system settings from a single interface. They can use the following commands to set and view the processor frequency scaling method:

`powersave -f` (processor runs at its fastest speed)
`powersave -l` (processor runs at its slowest speed)
`powersave -A` (processor speed is adjusted dynamically)
`powersave -c` (current setting is displayed)

Administrators can monitor the current processor speed using the command `/proc/cpuinfo`; for examples of the resulting information from this command, see the supplemental online section of this article at www.dell.com/powersolutions.

In the i386 SMP kernel of Red Hat Enterprise Linux 4, the MHz setting reported in /proc/cpuinfo does not update when the processor frequency changes, but the bogomips setting does. The MHz setting does update in the x86-64 uniprocessor kernel.

Demand-based power management not enabled in the BIOS

```
powernow-k8: Found 2 AMD Athlon 64 / Opteron processors (version 1.60.0)
powernow-k8: MP systems not supported by PSB BIOS structure
powernow-k8: init not cpu 0
```

Demand-based power management enabled in the BIOS

```
powernow-k8: Found 2 AMD Athlon 64 / Opteron processors (version 1.60.0)
powernow-k8:    0 : fid 0xa (1800 MHz), vid 0xa (1300 mV)
powernow-k8:    1 : fid 0x2 (1000 MHz), vid 0x12 (1100 mV)
powernow-k8: cpu_init done, current fid 0xa, vid 0xa
powernow-k8:    0 : fid 0xa (1800 MHz), vid 0xa (1300 mV)
powernow-k8:    1 : fid 0x2 (1000 MHz), vid 0x12 (1100 mV)
powernow-k8: cpu_init done, current fid 0xa, vid 0xa
```

Figure 4. Example output after loading the powernow-k8 driver in Red Hat Enterprise Linux 4 Update 4

Validating effective power management

Administrators can validate the effectiveness of AMD PowerNow! technology in a number of ways. For example, they can use AMD tools such as PSTCheck or System Stress Test to verify power state transitions and drops in processor clocking speed. Because other system components besides the processor also consume power, administrators should measure differences in power consumption based on the direct power supplied from the power socket.

Implementing efficient power management

Effective power management can play a key role in managing power use and controlling costs in enterprise data centers. By taking advantage of AMD PowerNow! technology on Dell PowerEdge servers with AMD Opteron processors, enterprises can enhance the efficiency of their power usage and thereby help reduce overall data center costs.

Manoj Gujarathi is a Dell OpenManage development manager, and has recently worked as software development manager for Dell PowerEdge servers with AMD processors. Manoj has an M.S. in Engineering from Washington State University and an M.S. in Computer Science from Texas Tech University.

Lisa Garrison is a senior manager in the Dell Enterprise Windows Engineering group working on integrating Microsoft Windows Server operating systems with server hardware. Lisa has a B.S. in Mathematics, a B.S. in Computer Science, and an M.S. in Business from Southeastern Oklahoma State University.

Jordan Hargrave is a senior software engineer in the Dell Enterprise Linux Engineering group. He has a strong background in dual-core processors and systems management. Jordan has a B.S. in Computer Science from Carnegie Mellon University and is currently pursuing a master's degree in Software Engineering from the University of Texas.

FOR MORE INFORMATION

Dell and AMD:
www.dell.com/amd

AMD power management:
enterprise.amd.com/us-en/AMD-Business/Technology-Home/Power-Management.aspx

Kerby, Brent. "Managing Data Center Power and Cooling with AMD Opteron Processors and AMD PowerNow! Technology." *Dell Power Solutions,* February 2007. www.dell.com/downloads/global/power/ps1q07-20070204-AMD.pdf

Who's Abusing Your Network?

Secure Networks™ for Higher Education

Today's open learning environment poses many security risks: students downloading music illegally, guest faculty connecting virus-infected laptops, or illegal/offensive content being hosted on your servers without your knowledge.

Secure Networks solutions from Enterasys address all of these challenges. Our security architecture delivers centralized command and control, enabling you to automatically enforce acceptable usage policies for users and applications. If a breach does occur, automated security response helps identify, locate and eliminate threats in your IT infrastructure.

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

Take the Secure Networks Challenge

See how our unique approach can secure any network from any vendor by scheduling a Secure Networks demo. Call **877-801-7082** or visit **enterasys.com/higher-ed**

enterasys®
Secure Networks™

There is nothing more important than our customers.

Performance Scaling

with Dell PowerEdge 2950 Servers and VMware Virtual Infrastructure 3

To assess virtualization scalability and performance, Dell engineers tested two Dell™ PowerEdge™ 2950 servers with dual-core Intel® Xeon® 5160 and quad-core Intel Xeon X5355 processors alongside a previous-generation PowerEdge 2850 server with dual-core Intel Xeon processors in a virtualized environment based on VMware® Virtual Infrastructure 3. The results show that the PowerEdge 2950 servers can provide enhanced performance and host more virtual machines than the PowerEdge 2850.

BY TODD MUIRHEAD

Related Categories:

Dell PowerEdge servers

Dell Enterprise Technology Center

Intel

Scalable enterprise

Virtualization

VMware

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

When deciding how to implement a virtualized environment with industry-standard components such as Dell PowerEdge servers and VMware Virtual Infrastructure 3 software, two key factors are how many virtual machines (VMs) a single server can run and what level of performance these VMs can provide. Although the answer depends on many different factors, testing with common applications and workloads can provide a starting point for enterprises planning such an environment.

In September and October 2006, Dell engineers tested virtualized environments based on VMware Virtual Infrastructure 3 on two differently configured PowerEdge 2950 servers and one previous-generation PowerEdge 2850 server. Using three different workloads, the test team assessed the number of VMs that each server could host and the overall system performance, and found that the PowerEdge 2950 servers offered significant gains in both regards compared to the PowerEdge 2850. These tests are intended to provide an example of the scalability and performance increases possible when migrating to a new enterprise IT environment.

Hardware configuration for test environment

The PowerEdge 2950 is a two-socket server that supports Intel Xeon 5000, 5100, and 5300 series processors. The Dell test team configured one PowerEdge 2950 with two dual-core Intel Xeon 5160 processors at 3.0 GHz, and another with two quad-core Intel Xeon X5355 processors at 2.66 GHz. The Intel Xeon 5160 has a 4 MB processor cache that is shared between the two cores and a 1,333 MHz frontside bus. The Intel Xeon X5355 is essentially two dual-core Intel Xeon 5160 processors combined, and so has an 8 MB level 2 (L2) cache (4 MB shared by each set of two cores) and a 1,333 MHz frontside bus.

An additional benefit of Intel Xeon 5160 processors is lower power consumption compared to previous-generation Intel Xeon processors. A previous study found that ninth-generation PowerEdge servers with Intel Xeon 5100 series

processors use up to 25 percent less power than eighth-generation PowerEdge servers with dual-core Intel Xeon processors.[1] The combination of lower power consumption and higher performance over previous-generation processors can result in increased performance per watt.

PowerEdge 2950 servers support up to 32 GB of RAM when using 4 GB dual in-line memory modules (DIMMs), but in the Dell tests both servers were configured with 16 GB of RAM using 2 GB, 667 MHz fully buffered DIMMs.

The PowerEdge 2850 server was configured with two dual-core Intel Xeon processors at 2.8 GHz—the fastest supported processors for this server—with a dedicated 2 MB cache for each core, 800 MHz frontside bus, and 8 GB of RAM using 400 MHz double data rate 2 (DDR2) DIMMs. Figure 1 summarizes the server configuration in the test environment.

The PowerEdge 2950 servers were connected to a storage area network (SAN) with a dual-port QLogic 2462 PCI Express host bus adapter (HBA) and utilized storage on a Dell/EMC CX3-80 array with twenty 146 GB, 15,000 rpm disks. The PowerEdge 2850 was connected to the SAN with two QLogic 2340 PCI Extended HBAs and utilized storage on a Dell/EMC CX700 array with twenty 73 GB, 10,000 rpm disks. The three types of VMs—each running a different workload, as described in the "Test workloads: Microsoft SQL Server 2005, SUSE LAMP, and NetBench" section in this article—were spread across the 20 disks on each storage array. These disks were divided into four 5-disk (4 + 1) RAID-5 logical units (LUNs). The three types of VMs were evenly divided across the LUNs so that a quarter of each type were on each LUN. Figure 2 summarizes the storage configuration used in the test environment.

	PowerEdge 2850 server	PowerEdge 2950 servers	
Processors	Two dual-core Intel Xeon processors DP at 2.8 GHz with two 2 MB L2 caches (one per core)	Two dual-core Intel Xeon 5160 processors at 3.0 GHz with 4 MB cache (shared)	Two quad-core Intel Xeon X5355 processors at 2.66 GHz with 8 MB cache (shared)
Frontside bus	800 MHz	1,333 MHz	
Memory	8 GB (400 MHz DDR2 DIMMs)	16 GB (667 MHz fully buffered DIMMs)	
Internal disks	Two Ultra320 SCSI 73 GB, 10,000 rpm drives	Two Serial Attached SCSI (SAS) 146 GB, 15,000 rpm drives	
Network interface cards (NICs)	Two 10/100/1,000 Mbps internal NICs and one Intel PRO/1000 XT Gigabit* Ethernet NIC	Two 10/100/1,000 Mbps internal NICs	
Disk controller	PowerEdge Expandable RAID Controller (PERC) 4/ei	PERC 5/i	
Fibre Channel HBA	Two QLogic 2340 PCI Extended HBAs	One QLogic 2462 PCI Express HBA	
Virtualization software	VMware ESX Server 2.5.2	VMware ESX Server 3	

* 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. Configurations for the PowerEdge 2850 server and two PowerEdge 2950 servers used in the test environment

	Dell/EMC CX700	Dell/EMC CX3-80
Attached servers	One PowerEdge 2850 server	Two PowerEdge 2950 servers
Controller cache	3,614 MB (2,818 MB write, 796 MB read)	10,384 MB (3,072 MB write, 7,312 MB read)
Fibre Channel speed	Fibre Channel 2 (FC2)	Fibre Channel 4 (FC4)
Disk enclosures	Two DAE2 disk array enclosures	Four DAE3P disk array enclosures
Disks	Twenty 73 GB, 10,000 rpm disks	Twenty 146 GB, 15,000 rpm disks
LUNs	Four 5-disk RAID-5 LUNs	Four 5-disk RAID-5 LUNs
Software	EMC® Navisphere® Manager and Access Logix™ software	EMC Navisphere Manager and Access Logix software

Figure 2. Configurations for the Dell/EMC CX700 and CX3-80 storage used in the test environment

Virtualization platform for test environment

The Dell tests used VMware Virtual Infrastructure 3 as the virtualization platform; this package includes ESX Server 3 and VirtualCenter 2 as well as features such as load balancing and VMware High Availability (VMware HA). ESX Server allows multiple VMs to run simultaneously on a single physical server. Each VM runs its own OS, which in turn has its own set of applications and services. Because ESX Server isolates each VM from other VMs on the

[1] Based on testing performed by Dell Labs in May 2006 using the SPECjbb2005 benchmark on a PowerEdge 2950 server with two dual-core Intel Xeon 5160 processors at 3.0 GHz (Woodcrest) and then with two dual-core Intel Xeon 5080 processors at 3.73 GHz (Dempsey); 4 GB of 667 MHz and 533 MHz fully buffered DIMM memory; two SAS 73 GB, 15,000 rpm hard disk drives; and the Microsoft Windows Server 2003 Enterprise x64 Edition OS, as compared to a PowerEdge 2850 server with two dual-core Intel Xeon processors at 2.8 GHz (Paxville); 4 GB of 400 MHz DDR2 memory; two SCSI 36 GB, 15,000 rpm hard disk drives; and Windows Server 2003 Enterprise x64 Edition OS with Service Pack 1 (SP1). Actual performance and power consumption will vary based on configuration, usage, and manufacturing variability. For more information about ninth-generation PowerEdge server performance, see "Scaling Business Applications with New Servers and Storage," by Mark Nickerson, Joe Pollock, Stori Waugh, and Stacy Hower, *Dell Power Solutions,* August 2006, www.dell.com/downloads/global/power/ps3q06-20060283-CoverStory.pdf; and "Improved Virtualization Performance with 9th Generation Servers," by David J. Morse, Dell Inc., August 2006, www.dell.com/downloads/global/solutions/9g%20_servers.pdf.

Figure 3. Processor utilization graph in VMware VirtualCenter 2 during a SUSE LAMP test

same physical server just as physical systems are isolated from one another, administrators have flexibility in using ESX Server to run different types of applications and operating systems at the same time. VirtualCenter 2 enables administrators to consolidate control and configuration of ESX Server systems and VMs, which can improve management efficiency in large environments.

The PowerEdge 2950 servers used ESX Server 3, and the PowerEdge 2850 server used ESX Server 2.5.2; both were managed by an existing VirtualCenter 2 console. All VMs were first tested on the PowerEdge 2850, then migrated to each PowerEdge 2950 in turn. The test team used VirtualCenter 2 to shut down the VMs, migrate them to a PowerEdge 2950, upgrade the virtual hardware and VMtools, and then restart the VMs.

Test workloads: Microsoft SQL Server 2005, SUSE LAMP, and NetBench

To compare the relative performance of the PowerEdge 2850 and PowerEdge 2950 servers, the test team ran three workloads on each server: the Microsoft® SQL Server™ 2005 database platform with an online transaction processing (OLTP) workload, the Novell® SUSE® Linux® Enterprise Server OS with a LAMP (Linux, Apache, MySQL, PHP) stack, and the Microsoft Windows Server® 2003 OS with NetBench 7.03.[2] To simulate how enterprises typically run applications on VMs using ESX Server in a production environment, the test team increased the number of VMs until processor utilization for the entire physical server exceeded 85 percent, with the highest utilization reaching 88 percent—a reasonably high level of usage that still allows for workload spikes. The test team calculated utilization levels by averaging the values from the esxtop utility run on the ESX Server service console during each test, and also monitored the utilization from the VirtualCenter console, as shown in Figure 3.

Each workload ran simultaneously on multiple VMs under the same load. By keeping all settings on the VM and driver systems identical and then observing how many VMs could be run simultaneously, the test team was able to measure how many VMs each physical server could support. Figure 4 shows the configuration for each type of VM in the test environment.

Microsoft SQL Server 2005. On the SQL Server 2005 VMs, the test team installed 32-bit versions of Microsoft Windows Server 2003 Release 2 (R2) Enterprise Edition and SQL Server 2005 with Service Pack 1 (SP1).[3] The SQL Server version of the Dell DVD Store database was loaded into SQL Server 2005 using the scripts provided with the DVD Store download to create the medium-size database. The complete DVD Store application code, including SQL Server and LAMP versions, is freely available for public use under the GNU General Public License (GPL) at linux.dell.com/dvdstore. The DVD Store database simulates the database back end of a simple Web-based storefront. The database size is small (approximately 1 GB), and representative of a database used for development or testing.

To simulate a load on the VMs, the test team used the DVD Store driver program, which is included in the DVD Store download. Each SQL Server 2005 VM was driven by four threads of the driver application with a 20-millisecond delay.

SUSE LAMP. For the LAMP workload, the test team installed 32-bit versions of Novell SUSE Linux Enterprise Server 9, Apache 2, and MySQL 5 on a VM. The MySQL version of the DVD Store application was loaded into MySQL 5, and the PHP version of the DVD Store application was set up on Apache. In this setup, the

Workload	Memory	Disk	Virtual NIC type	Number of virtual processors
Microsoft SQL Server 2005	512 MB	10 GB	Vmxnet	1
SUSE LAMP	1,024 MB	10 GB	Vlance/Flexible*	1
NetBench	512 MB	10 GB	Vmxnet	1

*The SUSE LAMP VMs used a Vlance NIC under ESX Server 2.5.2, but this was upgraded to Flexible as part of the migration to ESX Server 3.

Figure 4. Configurations for the virtual machines used in the test environment

[2] The SUSE LAMP and NetBench workloads used for these tests were also used in previous studies; for more information, see "VMware ESX Server Multiple Workload Performance on Dell PowerEdge 2850 and PowerEdge 6850 Servers," by Todd Muirhead, Dave Jaffe, and Scott Stanford, Dell Enterprise Product Group, July 2005, www.dell.com/downloads/global/solutions/vmware_mixed_workload.pdf; and "VMware ESX Server Performance Gains on Dell PowerEdge 2850 Dual Core Servers," by Todd Muirhead and Dave Jaffe, Dell Enterprise Product Group, November 2005, www.dell.com/downloads/global/solutions/esx_2850_dualcore.pdf. These two studies also used a SQL Server 2000 workload that was updated to use SQL Server 2005 for the tests described in this article.

[3] The use of Microsoft SQL Server 2005 in these tests does not indicate that Dell or Microsoft has tested or certified SQL Server with VMware virtualization software. As described at support.microsoft.com/?kbid=897615, Microsoft typically does not support problems with Microsoft operating systems or applications that run on VMs using non-Microsoft virtualization software unless the same problem can be reproduced outside the VM environment.

FINDING EXPERT ADVICE ON ENTERPRISE SOLUTIONS JUST GOT EASIER.

INTRODUCING THE DELL™ ENTERPRISE TECHNOLOGY CENTER.

It's your one-stop shop for expert advice on x86 standards-based enterprise solutions. Join the tech center community and get the latest tips, tools, and best-practice articles from experienced people you trust. Whether you need concrete technical strategies or just want to bounce around ideas, the tech center is designed with one organization in mind. Yours.

Join your peers today at www.dell.com/techcenter.

See the Dell difference at
www.dell.com/techcenter

DELL Purely You

	PowerEdge 2850 with dual-core Intel Xeon processors		PowerEdge 2950 with dual-core Intel Xeon 5160 processors			PowerEdge 2950 with quad-core Intel Xeon X5355 processors			
Workload	**Number of VMs**	**Performance**	**Number of VMs**	**Performance**	**Performance gain compared to PowerEdge 2850**	**Number of VMs**	**Performance**	**Performance gain compared to PowerEdge 2850**	**Performance gain compared to PowerEdge 2950 with Intel Xeon 5160 processors**
Microsoft SQL Server 2005	14	11,782 OPM	20	20,026 OPM	70%	32	29,346 OPM	149%	47%
SUSE LAMP	16	2,650 OPM	25	6,057 OPM	129%	44	9,852 OPM	272%	63%
NetBench	18	411 MB/sec	23	593 MB/sec	44%	42	1,001 MB/sec	144%	69%

Figure 5. Workload performance results for each server in the test environment

Web tier and the database tier ran on the same VM to create a complete LAMP stack.[4]

The driver for the LAMP stack differs from the driver used in the SQL Server testing in that it sends HTTP requests and receives HTML code returned from the Apache/PHP layer, whereas the SQL Server driver communicates directly with the database. However, the LAMP workload measures the same parameters: total orders per minute (OPM) handled by the application, and average response time experienced by the simulated customers. Each SUSE LAMP VM was driven by a single thread of the driver program with a 20-millisecond delay.

NetBench. NetBench 7.03, developed by *PC Magazine*, is a benchmark tool designed to simulate a file server workload. The program creates and accesses a set of files according to predefined scripts. NetBench is typically run with an increasing number of client engines running against a single server to measure how much throughput (in megabytes per second) can be achieved with a given number of connections.

The NetBench VMs were installed with the 32-bit version of Microsoft Windows Server 2003 R2 Enterprise Edition. To determine how many VMs could run on an ESX Server host, the test team increased the number of VMs and the number of client engines at the same rate until the processor utilization on the ESX Server host reached 85 percent. NetBench 7.03, with the included standard DiskMix script, was used with a 0.6-second think time to connect two client engines to each VM.[5] This simulates multiple file servers on the same ESX Server host, similar to a file server consolidation scenario. The driver systems on which the client engines ran had mapped drives to all the test VMs. In NetBench the test directories path file was modified so that as successive client engines were added, they would use the next drive letter, which corresponded to the next VM.

Test results measuring scalability and performance

The test team first ran the VMs on the PowerEdge 2850 server in successive tests, adding VMs in each round as described in the "Test workloads: Microsoft SQL Server 2005, SUSE LAMP, and NetBench" section in this article. Next, they migrated the VMs to the PowerEdge 2950 server with Intel Xeon 5160 processors and repeated the tests. Finally, they migrated the VMs to the PowerEdge 2950 server with Intel Xeon X5355 processors and repeated the tests once again.

The difference in the number of VMs and the associated performance metric—OPM for SQL Server 2005 and SUSE LAMP and megabytes per second for NetBench—indicated the relative difference in performance. The test team calculated the performance results for the SQL Server 2005 and SUSE LAMP VMs by totaling the OPM from all the VMs running in the test environment; NetBench provides the megabytes-per-second metric as part of the results displayed at the end of a test. Figure 5 summarizes the results for the three workloads on each server.

For all three workloads, the number of VMs that each PowerEdge 2950 server could host was consistently higher than the number the PowerEdge 2850 could host, with the largest number of VMs achieved by the PowerEdge 2950 server with quad-core Intel Xeon X5355 processors running the SUSE LAMP workload. This PowerEdge 2950 server, running with eight cores and 16 GB of RAM, was able to host more than 40 VMs for two of the three workloads.

The PowerEdge 2950 servers also showed significant gains in workload performance when compared with the PowerEdge 2850. The PowerEdge 2950 server with Intel Xeon X5355 processors showed much larger performance gains than the PowerEdge 2950 with Intel Xeon 5160 processors, with the largest gain—272 percent—generated by the SUSE LAMP workload. Figure 6 illustrates the relative

[4] The LAMP stack has been fully documented in "MySQL Network and the Dell PowerEdge 2800: Capacity Sizing and Performance Tuning Guide for Transactional Applications," by Todd Muirhead, Dave Jaffe, and Nicolas Pujol, Dell Enterprise Product Group, April 2005, www.dell.com/downloads/global/solutions/mysql_network_2800.pdf.

[5] The NetBench client driver systems were two Dell PowerEdge 6650 servers with four Intel Xeon processors at 2.8 GHz, 8 GB of RAM, Intel Gigabit Ethernet adapters, and Windows Server 2003. The NetBench client driver systems and ESX Server hosts were connected to a Dell PowerConnect™ 5224 Gigabit Ethernet switch. The NetBench controller ran Windows Server 2003 Enterprise Edition.

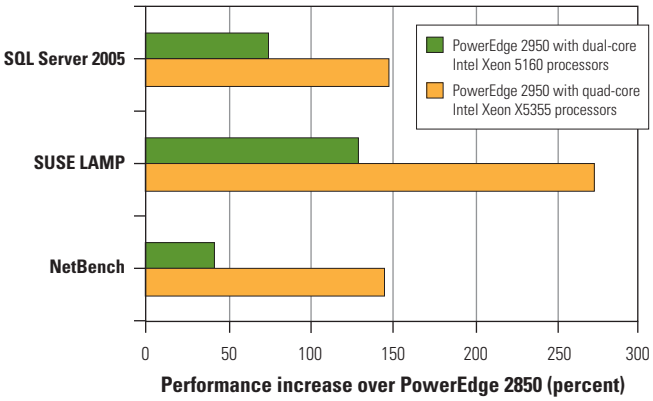


Figure 6. Relative performance increases for the PowerEdge 2950 servers over a PowerEdge 2850 server with dual-core Intel Xeon processors

performance increases for the PowerEdge 2950 servers over the PowerEdge 2850 server.

Enhanced virtualization performance

Dual-core Intel Xeon 5160 and quad-core Intel Xeon X5355 processors are designed to provide a significant boost in performance over previous-generation dual-core Intel Xeon processors. Performance increases may vary depending on the workload and other factors, but in the Dell tests described in this article, the PowerEdge 2950 performance gains were significant enough to allow more VMs to be hosted, with greater overall system throughput, than was possible on the PowerEdge 2850—including performance gains of up to 272 percent on a PowerEdge 2950 server with quad-core Intel Xeon X5355 processors.

The performance gains of the quad-core Intel Xeon X5355 over the dual-core Intel Xeon 5160 demonstrate that in virtualized environments, additional cores enable the hosting of additional VMs and help increase overall performance. These results suggest the type of enhanced scalability and performance possible with ninth-generation Dell PowerEdge servers and VMware virtualization software, which enterprises can take into account when planning to implement or upgrade a virtualized environment.

Todd Muirhead is a senior engineering consultant on the Dell Enterprise Technology Center team. Todd has a B.A. in Computer Science from the University of North Texas and is Microsoft Certified Systems Engineer + Internet (MCSE+I) certified.

TALK BACK

Tell us how the Dell Enterprise Technology Center can help your organization better simplify, utilize, and scale enterprise solutions and platforms. Send your feedback and ideas to **setc@dell.com**.

Comparing Enterprise Database Performance

on Dell PowerEdge and Sun Fire Servers

To evaluate the performance of dual-core processor–based servers from Dell and Sun, Dell engineers used the open source Dell DVD Store database to test a two-node Dell™ PowerEdge™ 2950 server cluster and a PowerEdge 6850 server with Intel® Xeon® processors alongside a similarly configured Sun Fire V490 server with Sun UltraSPARC IV+ processors. Not only did the Dell servers handle more orders per minute than the Sun server, but both also cost significantly less per order handled.

BY DAVE JAFFE, PH.D., AND TODD MUIRHEAD

Related Categories:

Characterization

Dell PowerEdge servers

Dell Enterprise Technology Center

Dell/EMC storage

Microsoft Windows Server 2003 x64 Editions

Performance

Red Hat Enterprise Linux

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

Several major server processor vendors have recently introduced processors with multiple 64-bit cores and large caches in the same module. The fastest such processor from Intel for two-socket systems is the 64-bit dual-core Intel Xeon 5160 processor at 3.0 GHz with a 4 MB level 2 (L2) cache and a 1,333 MHz frontside bus, and the fastest for four-socket systems is the 64-bit dual-core Intel Xeon 7140 processor MP at 3.4 GHz with a 16 MB level 3 (L3) cache. The fastest such processor from Sun for four-socket systems is the dual-core Sun UltraSPARC IV + processor at 1.8 GHz with a 32 MB L3 cache. Sun claims that its servers featuring these processors are more than five times faster than previous-generation servers.[1]

To compare the performance of Dell and Sun servers that feature these dual-core processors, in October 2006 Dell engineers used an online transaction processing (OLTP) workload with the open source Dell DVD Store database to measure the relative performance of a two-node Dell PowerEdge 2950 cluster with two dual-core Intel Xeon 5160 processors at 3.0 GHz per server, a PowerEdge 6850 server with four dual-core Intel Xeon 7140 processors MP at 3.4 GHz, and a Sun Fire V490 server with four dual-core UltraSPARC IV + processors at 1.8 GHz. The results showed that the Dell servers as configured outperformed the Sun server, and did so with a lower overall cost and lower price/performance.

Hardware test configuration

The test team compared servers configured with the most powerful dual-core processors available for each platform.

[1] "Sun Fire V490 Server – At a Glance," by Sun Microsystems, Inc., www.sun.com/servers/midrange/v490/index.xml.

	Sun Fire V490	Each Dell PowerEdge 2950	Dell PowerEdge 6850
OS	Sun Solaris 10 (1/06)	64-bit Red Hat Enterprise Linux AS 4	Microsoft Windows Server 2003 R2 Enterprise x64 Edition
Processors	Four dual-core Sun UltraSPARC IV+ processors at 1.8 GHz with 32 MB L3 cache	Two dual-core Intel Xeon 5160 processors at 3.0 GHz with 4 MB L2 cache	Four dual-core Intel Xeon 7140 processors MP at 3.4 GHz with 16 MB L3 cache
Memory	16 GB (thirty-two 512 MB dual in-line memory modules [DIMMs])	8 GB (four 2 GB, 667 MHz fully buffered DIMMs)	16 GB (eight 2 GB double data rate 2 [DDR2] DIMMs)
Internal disks	Two Fibre Channel–Arbitrated Loop (FC-AL) 146 GB, 15,000 rpm drives	Two Serial Attached SCSI (SAS) 146 GB, 15,000 rpm drives	Two Ultra320 SCSI 146 GB, 15,000 rpm drives
Network interface cards (NICs)	Two internal 10/100/1,000 Mbps NICs	Two internal and two external 10/100/1,000 Mbps NICs	Two internal 10/100/1,000 Mbps NICs
Disk controller	FC-AL (embedded)	PowerEdge Expandable RAID Controller (PERC) 5/i (embedded)	LSI Logic Ultra320 SCSI (embedded)
Fibre Channel HBAs	Two QLA2340 Fibre Channel 2 (FC2) PCI-X HBAs	One dual-port QLogic QLE2462 Fibre Channel 4 (FC4) PCI Express HBA	One dual-port QLogic QLE2462 FC4 PCI Express HBA
Remote management	System controller card	Dell Remote Access Controller 5 (DRAC 5), without modem	DRAC 4, without modem
Hardware price as configured with three-year Gold-level support*	US$91,557	US$26,852 for two-node cluster ($13,426 per server)	US$25,747
Database Enterprise Edition license price	US$240,000 (8 cores × 0.75 × $40,000 for database software)	US$240,000 (8 cores × 0.5 x [$40,000 for database software + $20,000 for clustering software])	US$160,000 (8 cores × 0.5 × $40,000 for database software)
Total price (hardware plus database)	US$331,557	US$266,852	US$185,747

*Hardware prices are given as listed at www.sun.com and www.dell.com on October 12, 2006.

Figure 1. Configurations and pricing for the Sun Fire V490, Dell PowerEdge 2950, and Dell PowerEdge 6850 servers used in the test environment

Each PowerEdge 2950 was configured with two dual-core Intel Xeon 5160 processors at 3.0 GHz with a 4 MB L2 cache shared between the two cores. The PowerEdge 6850 was configured with four dual-core Intel Xeon 7140 processors MP at 3.4 GHz with a 16 MB L3 cache shared between the two cores. The Sun Fire V490 was configured with four dual-core UltraSPARC IV+ processors at 1.8 GHz with a 32 MB L3 cache. The two stand-alone systems (the PowerEdge 6850 and Sun Fire V490) had 16 GB of RAM, two internal disks, and remote access capabilities. Each PowerEdge 2950 had 8 GB of RAM, plus two extra network ports to provide a redundant cluster interconnect. Thus both stand-alone systems as well as the cluster had a total of 16 GB of RAM and eight processor cores. The PowerEdge 2950 cluster ran the 64-bit version of the Red Hat® Enterprise Linux® AS 4 OS, the PowerEdge 6850 ran the Microsoft® Windows Server® 2003 Release 2 (R2) Enterprise x64 Edition OS, and the Sun Fire V490 ran the Solaris 10 (1/06) OS, which is also 64-bit.

Figure 1 summarizes the server configuration in the test environment along with the pricing of each configuration, not including storage area network (SAN) hardware or software. The price of the database software is different for Dell and Sun systems because of the different modifiers for each architecture type: each Intel Xeon core is charged at 50 percent, while each UltraSPARC core is charged at 75 percent.

All three systems used a SAN-attached Dell/EMC CX3-80 Fibre Channel storage array. Each server was attached to the SAN through QLogic host bus adapters (HBAs). The two PowerEdge models support PCI Express, so they used QLogic QLE2462 HBAs. The Sun Fire V490 only supports PCI Extended (PCI-X), so it used QLogic QLA2340 HBAs.

Each server was assigned to a set of storage logical units (LUNs) that used the same number and type of disk drives. The test team followed the guidelines in the EMC® Support Matrix, which includes both PowerEdge servers and the Sun Fire V490,

TALK BACK

Tell us how the Dell Enterprise Technology Center can help your organization better simplify, utilize, and scale enterprise solutions and platforms. Send your feedback and ideas to **setc@dell.com**.

to help ensure that they used the proper driver and firmware versions.[2] The PowerEdge 2950 servers used the 8.01.02-d4 version of the QLogic QLE2462 driver for Linux and EMC PowerPath® 4.5.1 software for Linux. The PowerEdge 6850 used the 9.1.2.15 version of the QLogic QLE2462 driver for the Microsoft Windows® OS and PowerPath 4.5.0 for Windows. The Sun Fire V490 used the QLogic FCode version 2.00.09 on the QLogic QLA2340 HBAs and PowerPath 5.0.0. The PowerPath software provided load balancing and failover for the dual HBAs. Figure 2 summarizes the storage components used in the test environment.

Controller	Dell/EMC CX3-80
Disk enclosures	Four Dell/EMC DAE3P enclosures
Disks	Fifty 73 GB, 15,000 rpm drives
LUNs	• Four 10-Disk RAID-10 LUNs (for data) • Two 2-Disk RAID-1 LUNs (for logs) • One 5-Disk RAID-0 LUN (for temporary data staging for loading) • One hot-spare disk
Software	EMC Navisphere® Manager, Access Logix™, and PowerPath software

Figure 2. Storage components used for the Sun Fire V490, Dell PowerEdge 2950, and Dell PowerEdge 6850 servers in the test environment

Database test configuration

The test team installed and configured a leading enterprise database platform on the PowerEdge 2950, PowerEdge 6850, and Sun Fire V490 servers according to the vendor's installation guide for each system, and applied the most recent patch sets. The test team then used the platform's database creation assistant wizard to create the database instances on each system, and used the Dell DVD Store build scripts to create and load the test database. The large version of the database was used, which consists of approximately 100 GB of data, indexes, and stored procedures.

The test team used the database platform auto-tuning features to configure the database parameters. On each PowerEdge 2950 server, approximately 6 GB of the 8 GB of RAM was assigned to the database, and the database platform automatically configured how this memory would be allocated. On the PowerEdge 6850 and Sun Fire V490 servers, approximately 12 GB of the 16 GB of RAM was assigned to the database and allocated similarly. Figure 3 lists the resulting database initialization parameters.

All three systems used the database platform's automatic storage manager for the database data files and log files. This storage manager helps reduce file system overhead by allowing the database to access and manage the storage more directly than would otherwise be the case.

The test team used the same database tablespace configuration on each system, as shown in Figure 4. The four RAID-10 LUNs were put into one disk group in the storage manager, which was used for the data, index, undo, and temporary tablespaces. In a similar fashion, the two RAID-1 LUNs were assigned to a disk group and used for redo logs.

Database test workload

To measure system performance, the test team used the large (100 GB) version of the DVD Store database, which represents an online DVD store carrying 1 million titles. The complete DVD Store application code is freely available for public use under the GNU General Public License (GPL) at linux.dell.com/dvdstore.

The test team used the included driver programs to simulate orders, and tracked orders per minute (OPM) as the performance metric. The tests used the DVD Store Version 2 (DS2) database test application, which includes advanced database features such as transactions, triggers, and referential integrity constraints along with functionality typical of some online stores, such as reporting previous purchases and recommending titles to users.

The database was driven by a C# program simulating users logging in to the online store; browsing for DVDs by category, title, or actor; and then submitting an order. The driver program measures

Parameter	Sun Fire V490	Each Dell PowerEdge 2950	Dell PowerEdge 6850
DB_BLOCK_SIZE	8192	8192	8192
DB_CACHE_SIZE	11727273984	5553258496	11727273984
JAVA_POOL_SIZE	16777216	16777216	16777216
LARGE_POOL_SIZE	16777216	16777216	16777216
OPEN_CURSORS	600	600	600
SESSION_CACHED_CURSORS	40	40	40
PROCESSES	150	150	150
PGA_AGGREGATE_TARGET	418381824	408944640	418381824
SGA_MAX_SIZE	12582912000	6291456000	12582912000
SGA_TARGET	12582912000	1227882496	12582912000
SHARED_POOL_SIZE	805306368	687865856	805306368
UNDO_RETENTION	300	300	300

Figure 3. Database initialization parameters used in the test environment

[2] The EMC Support Matrix is available at www.emc.com/interoperability/index.jsp.

Server management on the fly.

Access the data center from the comfort of anywhere.

The Dell™ 2161DS-2/4161DS remote console switches give today's server administrators the power to monitor and control servers from virtually anywhere. For maximum flexibility, the 4161DS provides **access for up to four simultaneous users and by 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, across the globe, or in yet another airport, your data center can always be within your reach.**

Digital Availability. Easy as DELL

www.dell.com

the OPM that the database can handle as well as the total response time experienced by the simulated users.

Database schema

The DVD Store database comprises eight tables, as shown in Figure 5. The Customers table is pre-populated with 200 million customers—100 million customers from the United States and 100 million customers from the rest of the world. The Orders table is pre-populated with 10 million orders per month for a full year. The Orderlines and Cust_Hist tables are pre-populated with an average of five items per order. The Products table contains 1 million DVD titles. To help improve performance, the quantities in stock and sales figures per product are placed in a separate table, Inventory. When the QUAN_IN_STOCK value for a product in the Inventory table falls below a specified value, a trigger is activated to write information about the product to the Reorder table. In a real-world implementation, a separate process (not modeled) would watch this table to initiate reordering of needed titles. Finally, the Categories table listed the 16 DVD categories.

Tablespace	Contents	Space used/available
CUSTTBS	Customers table	50 GB/51 GB
INDXTBS	Indexes	40 GB/40 GB
ORDERTBS	Orders and Orderlines tables	25 GB/30 GB
DS_MISC	Products and Categories tables	0.03 GB/0.5 GB
UNDOTBS	Undo tablespace	0.5 GB/2 GB
TEMP	Temporary table	1 GB/3 GB

Figure 4. Database tablespaces used in the test environment

Table	Columns	Number of rows
Customers	CUSTOMERID, FIRSTNAME, LASTNAME, ADDRESS1, ADDRESS2, CITY, STATE, ZIP, COUNTRY, REGION, EMAIL, PHONE, CREDITCARDTYPE, CREDITCARD, CREDITCARDEXPIRATION, USERNAME, PASSWORD, AGE, INCOME, GENDER	200 million
Orders	ORDERID, ORDERDATE, CUSTOMERID, NETAMOUNT, TAX, TOTALAMOUNT	120 million
Orderlines	ORDERLINEID, ORDERID, PROD_ID, QUANTITY, ORDERDATE	600 million
Cust_Hist	CUSTOMERID, ORDERID, PROD_ID	600 million
Products	PROD_ID, CATEGORY, TITLE, ACTOR, PRICE, SPECIAL, COMMON_PROD_ID	1 million
Inventory	PROD_ID, QUAN_IN_STOCK, SALES	1 million
Reorder	PROD_ID, DATE_LOW, QUAN_LOW, DATE_REORDERED, QUAN_REORDERED, DATE_EXPECTED	Variable
Categories	CATEGORY, CATEGORYNAME	16

Figure 5. Dell DVD Store database schema

Database stored procedures

The DVD Store database is managed through six stored procedures (Login, New_customer, Browse_by_category, Browse_by_title, Browse_by_actor, and Purchase) and one trigger (Restock). The first two stored procedures are used during the login phase. If the customer is returning, Login retrieves the customer's information, in particular CUSTOMERID; if the customer is new, New_customer creates a new row in the Customers table with the user data. Following login, customers search for a DVD by category, title, or actor, which are implemented by Browse_by_category, Browse_by_title, and Browse_by_actor, respectively. Finally, after customers make their selections, Purchase completes the transaction.

The stored procedures of the DS2 application include features to enhance modeling of online stores. During the Login procedure, for example, the customer's previous order (up to 10 titles) is reported, along with titles that other customers who like those titles have recommended. The Browse_by_category procedure returns titles in the specified category that are currently on sale. The Purchase procedure checks the QUAN_IN_STOCK field from the Inventory table to see whether a title is available; because this task is carried out using a database transaction, if there is insufficient quantity to fill the order, the QUAN_IN_STOCK data is not updated, and no new record is written to the Orders table. Finally, when the QUAN_IN_STOCK value of a particular title falls below 10, the Restock database trigger is activated, causing an entry to be written in the Reorder table.

Online transaction processing driver application

The test team used a multi-threaded driver program included with the DS2 application to simulate an OLTP workload. Each thread of the OLTP driver application connected to the database and made a series of stored procedure calls that simulated users logging in, browsing, and purchasing. An extremely small delay, or *think time,* was used so that the database connections remained full, simulating what happens in a real multitiered application in which a small number of connections are pooled and shared among Web servers that may be handling thousands of simultaneous customers. This approach allowed a realistic simulation of database activity without modeling thousands of users.

Each thread of the OLTP driver modeled a series of customers going through the entire sequence of logging in, browsing the catalog several ways, and finally purchasing the selected items. Each completed sequence by a customer counted as a single order. The driver measured order rates and the average response time to complete each order. Several configurable parameters were used to control the application, as described in Figure 6.

Test results: Performance and price/performance

The tests measured the OPM each system could handle while keeping processor utilization at approximately 85 percent, a reasonable level that still allowed for workload spikes. The test team measured processor utilization using mpstat in Red Hat Enterprise Linux, the Performance Monitor program in Windows Server 2003, and vmstat in Solaris.

Figure 7 summarizes the results. The two-node PowerEdge 2950 cluster handled 26 percent more OPM than the Sun Fire V490, and the PowerEdge 6850 handled 83 percent more OPM. Because the Sun Fire V490 also cost 24 percent and 78 percent more than the PowerEdge 2950 cluster and PowerEdge 6850 server, respectively, the Dell systems also offer lower price/performance (dollars/OPM)—the Sun Fire V490 cost 56 percent more per order handled than the PowerEdge 2950 cluster and 226 percent more per order handled than the PowerEdge 6850 server.

Parameter	Description	Values used in test
n_threads	Number of simultaneous database connections	• Sun Fire V490: 13 • Dell PowerEdge 2950: 16 • Dell PowerEdge 6850: 36
reset_time	Warm-up time before statistics are kept	2 minutes
run_time	Runtime during which statistics are kept	10 minutes
think_time	Amount of time a thread waits between orders	0 seconds
pct_returning	Percentage of users that are returning users	80%
pct_new	Percentage of users that are new users	20%
n_searches	Number of searches per order	• Range: 1–5 • Average: 3
search_batch_size	Number of items returned in each search	• Range: 1–9 • Average: 5
n_line_items	Number of items purchased	• Range: 1–9 • Average: 5

Figure 6. Dell DVD Store OLTP driver parameters used in the test environment

Enhanced performance and cost-effectiveness with Dell PowerEdge servers

In the Dell tests, the two-node Dell PowerEdge 2950 cluster with four dual-core Intel Xeon 5160 processors at 3.0 GHz achieved 26 percent higher performance compared with the Sun Fire V490 server with four dual-core Sun UltraSPARC IV+ processors at 1.8 GHz, and the Dell PowerEdge 6850 server with four dual-core Intel Xeon 7140 processors MP at 3.4 GHz achieved 83 percent higher performance compared with the Sun Fire V490. When considering an IT investment, however, enterprises should consider not only performance but also overall price and price/performance—and as configured in these tests, the Dell systems not only outperformed the Sun system, but also cost less. The Sun Fire V490 cost 24 percent more overall than the PowerEdge 2950 cluster and 78 percent more overall than the PowerEdge 6850 server, with the Dell servers providing price/performance advantages of 56 percent and 226 percent, respectively. These advantages illustrate the strong value industry-standard servers, software, and operating systems can provide in enterprise data centers compared with similar proprietary setups.

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

Todd Muirhead is a senior engineering consultant on the Dell Enterprise Technology Center team. Todd has a B.A. in Computer Science from the University of North Texas and is Microsoft Certified Systems Engineer + Internet (MCSE+I) certified.

System	Number of simultaneous database connections (see Figure 6)	Orders per minute (OPM; higher is better)	Average response time	Dell system performance advantage compared with Sun Fire V490	Total price (hardware plus database—see Figure 1)	Price/performance (dollars/OPM; lower is better)	Dell system price/performance advantage compared with Sun Fire V490
Sun Fire V490 server	13	17,660	0.038 seconds	—	US$331,557	US$18.77	—
Two-node Dell PowerEdge 2950 cluster	16	22,169	0.039 seconds	26%	US$266,852	US$12.04	56%
Dell PowerEdge 6850 server	36	32,264	0.059 seconds	83%	US$185,747	US$5.76	226%

Figure 7. Performance and price/performance test results

Introducing the Dell PowerVault RD1000:

A Portable Disk-based Replacement for Traditional Low-End Tape Backup

The new Dell™ PowerVault™ RD1000 removable disk drive is designed to cost-effectively match or surpass the advantages of traditional entry-level and low-end tape backup while providing the backup and restore performance of a hard disk drive. This article discusses the potential benefits of deploying the PowerVault RD1000 in enterprise data centers.

BY CURT KREMPIN

Related Categories:

Backup, recovery, and archiving (BURA)

Disaster recovery

Storage

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

Attempts to entirely replace traditional entry-level and low-end tape backup with different technologies have typically failed because they could not match tape's portability, capacity, archiving capability, or cost-effectiveness. The IT industry has therefore been exploring technologies that both complement tape and address its biggest disadvantages: its relatively low performance during backup and restore operations, and its relatively high failure rates compared with hard disk drives (HDDs).

One alternative has been disk-to-disk (D2D) systems. However, these systems are not typically deployed alone as a true tape replacement. Instead, rather than backing up data directly to tape, enterprises use D2D systems as an interim step to help increase performance: the backup application writes data to the D2D target, and then later moves the data to tape for off-site disaster recovery and long-term storage—an architecture known as disk-to-disk-to-tape (D2D2T). Although D2D2T has performance and availability advantages over a tape-only design, these advantages typically also increase system costs and complexity.

The new Dell PowerVault RD1000 is designed to replace traditional entry-level and low-end tape entirely while still maintaining the advantages of D2D systems (see Figure 1). The PowerVault RD1000 is a removable backup disk drive that cost-effectively matches and surpasses key advantages of tape and provides the backup and restore performance of random access HDDs. It can be used in many ways, but is primarily intended to back up critical data for high-end desktops and entry-level servers.

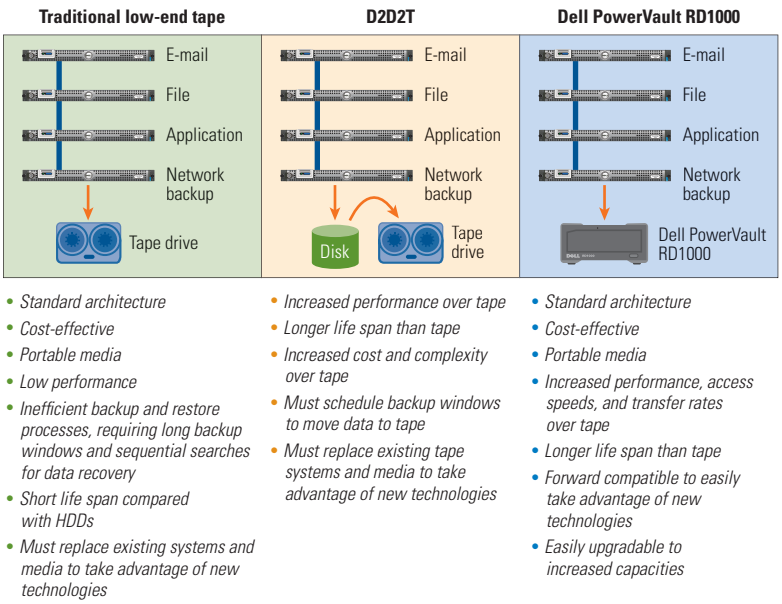


Figure 1. Backup using the Dell PowerVault RD1000 compared with traditional low-end tape and disk-to-disk-to-tape backup

Introducing the Dell PowerVault RD1000

The PowerVault RD1000 enables the same traditional backup process as tape—that is, moving data directly to a device with removable media. To a host computer or server, the PowerVault RD1000 removable disk cartridge looks like a tape cartridge. However, backup performance is significantly increased: the PowerVault RD1000 can take less than an hour to back up 80 GB of native data at its 25 MB/sec external USB transfer rate (up to 30 MB/sec internal Serial ATA [SATA]), while the same operation can take a typical digital linear tape (DLT) VS160 drive over 2 hours, a typical DAT72 drive over 6 hours, and a typical Travan drive over 11 hours (see Figure 2).

The 3.5-inch PowerVault RD1000 uses a removable 2.5-inch HDD suspended in a durable cartridge and designed for portability—the same drives typically used in notebook computers because of their size and locking-head feature. Small-form-factor HDDs like these have undergone significant advancements recently, resulting in enhanced mechanical reliability and life span, including a mean time to failure of 550,000 hours. Features such as ramp-load heads and fluid dynamic bearings are designed to eliminate head-media contact and disk sticking. Its protective design has enabled the removable disk cartridge to pass Dell drop tests of nearly 1 meter onto a tiled concrete floor without damage.

In addition, the PowerVault RD1000 is designed to be easy to deploy, avoiding complex setup procedures; to be compatible with most common backup applications; and to offer plug-and-play capability in most backup architectures. Deploying it should not typically require changing storage designs, complicating backup processes, or adding unnecessary costs.

Comparing the Dell PowerVault RD1000 with low-end tape

In addition to providing the advantages described in the preceding section, the PowerVault RD1000 is designed to match or exceed traditional low-end tape media in performance, reliability, cost-effectiveness, and simplicity of backup and restore operations.

Performance. Like tape drives, HDDs vary in throughput and performance. The advantage that HDDs have over tape is the ability to randomly access recorded data—even if data is written in a sequential format, the PowerVault RD1000 can access and read the data randomly, helping essentially eliminate seek time and significantly reducing single-file restore times. And because PowerVault RD1000 disk cartridges are available in native capacities ranging from 40 GB to 120 GB, this storage system's high capacity combined with its USB transfer speed of up to 25 MB/sec for the external version and SATA transfer speed of up to 30 MB/sec for the internal version enable enterprises to perform full backups every day in much less time than it takes to perform even incremental backups to low-end tape, which typically transfers at about 3 MB/sec.

Reliability. A tape drive's life span can be limited by the magnetic head, which makes physical contact with the tape media as it reads and writes data. This contact degrades both the head and the tape, and can eventually result in the failure of both the drive and the media. Because the PowerVault RD1000 has no such direct contact and features the same simple connector interface typical of HDDs, it has an expected reliability at least 10 times better than typical low-end tape drives, with its life span limited primarily by the connector rather than the cartridge itself. These physical advantages allow the PowerVault

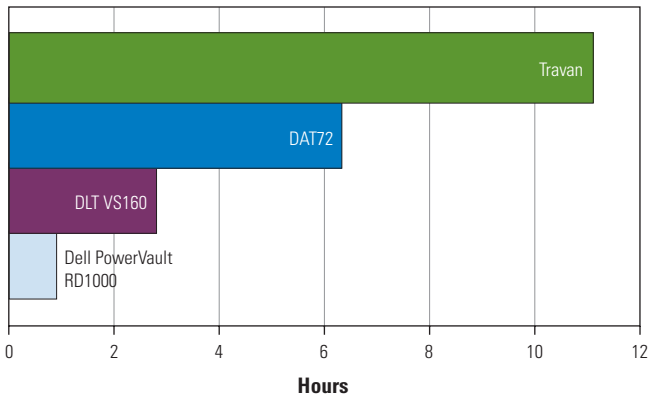


Figure 2. Typical time to back up 80 GB of native data for the Dell PowerVault RD1000 compared with common tape media

	Dell PowerVault RD1000	DAT72	Travan
Native capacity	40 GB to 120 GB (with planned expansion to 160 GB)	36 GB	20 GB
Native SATA performance	Up to 30 MB/sec (external USB performance up to 25 MB/sec)	3.5 MB/sec	2 MB/sec
Backward and forward compatibility	✔		
Electrostatic discharge (ESD)–protected cartridge	✔		
Drag-and-drop functionality	✔		
Cartridge status indicator	✔		
No required cartridge cleaning	✔		
Mean time to failure	550,000 hours	125,000 hours	370,000 hours
Specified media uses	5,000	200	200
Typical relative price	1*x*	2*x*	1.5*x*

Figure 3. Comparison of Dell PowerVault RD1000 removable disk drive with DAT72 and Travan tape drives

RD1000 cartridge to handle more than 5,000 load/unload actions, giving the media a usage life span more than 25 times longer than that of typical low-end tape media.

Cost-effectiveness. To take advantage of ongoing storage capacity advancements for tape drives, enterprises must regularly purchase new tape drives and the associated new media. The PowerVault RD1000, however, features both backward and forward compatibility, helping avoid obsolescence problems: as higher-capacity cartridges are introduced, they can work with existing docks, and vice versa. The simplicity of the PowerVault RD1000 design therefore allows for superior total cost of ownership when compared to low-end tape.

Simplicity of backup and restore operations. Using simple drag-and-drop functionality can make managing a PowerVault RD1000 with a backup application easier than using a low-end tape device. In addition, the PowerVault RD1000 dock automatically tracks how many times a cartridge has been loaded, saving administrators from the tedious task of manually tracking cartridge use—an especially valuable feature in complex media rotation systems.

Restoring an entire disk volume from incremental low-end tape backups, meanwhile, requires going through every piece of tape media that has been used in the backup process—and if the backup catalog is somehow contaminated, it can only be re-created by another time-consuming search of every piece of affected media. Even with an intact catalog, however, finding a file requires a slow serial search of the correct tape cartridge. The PowerVault RD1000, in contrast, uses the random access approach of HDDs to locate and retrieve data, helping significantly simplify file recovery compared with tape.

Figure 3 summarizes how the PowerVault RD1000 surpasses two common tape technologies, DAT72 and Travan, in performance, feature set, and cost-effectiveness.

Creating a single disk-based solution for backup and restore

Before the arrival of removable disk technology, enterprises had to choose between tape, disk, or a combination to back up their high-end desktops and low-end servers, with each method having its own architecture, performance, and cost trade-offs. The new Dell PowerVault RD1000 removable disk drive is designed to provide a viable alternative to these choices, combining the portability, capacity, archiving capability, and cost-effectiveness of tape with the high performance, simplicity, and reliability of disks in a single backup drive.

Curt Krempin is a product marketing manager in the Dell Enterprise Product Group. He is currently responsible for product marketing for tape drives and entry-level disk backup. Curt has a B.S. in Computer and Electrical Engineering from the University of Florida and an M.B.A. from the University of Texas.

FOR MORE INFORMATION

Dell storage:
www.dell.com/storage

Deploying the Dell PowerVault NX1950 for High-Availability Storage

Complex enterprise storage infrastructures can present numerous manageability, scalability, and availability challenges. The Dell™ PowerVault™ NX1950 storage server helps meet these challenges by providing flexible, scalable, and highly available network attached storage services designed to integrate seamlessly with existing networks and infrastructures.

BY KEVIN GUINN AND ANANDA SANKARAN

Related Categories:

Dell PowerVault storage

High availability (HA)

Network attached storage (NAS)

Storage

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

The Dell PowerVault NX1950 network attached storage solution helps simplify storage deployment and management as well as integration with existing infrastructures. There are two variants of this network attached storage solution: an integrated approach that employs the Dell PowerVault MD3000 Serial Attached SCSI (SAS) disk array, and a storage area network (SAN) gateway that is designed to integrate seamlessly with new or existing Dell/EMC SAN deployments. The integrated approach supports clusters of two PowerVault NX1950 head units, and the SAN gateway supports clusters with up to eight PowerVault NX1950 head units.

Understanding Dell PowerVault NX1950 services and interfaces

The Dell PowerVault NX1950 uses Microsoft® Windows® Unified Data Storage Server 2003 to provide storage services such as file serving and to serve as an Internet SCSI (iSCSI) target device. Its file-serving feature includes support for protocols such as Common Internet File System (CIFS), Network File System (NFS), and Distributed File System (DFS)—enabling clients using UNIX®, Mac, and Windows platforms to access data shares on the file server. The iSCSI target service enables the PowerVault NX1950 to function as an iSCSI-based storage device that can provide block-level access to logical disk storage; hosts capable of connecting to storage devices through iSCSI can connect to the PowerVault NX1950 over an IP network, and the iSCSI target service provides logical units (LUNs) to the initiator hosts. The exposed disks are internally managed in the storage server as virtual disks using the virtual hard disk (.vhd) file format. The iSCSI target also allows the creation of snapshots—point-in-time read-only copies—of these virtual disks to facilitate backup and recovery. A Microsoft Volume Shadow Copy Service (VSS) provider allows the snapshots to be created and used by the initiator hosts.

The PowerVault NX1950 solution includes simple, unified management capabilities. Administrators can carry out setup and configuration through the PowerVault NX1950

Configuration Tasks wizard. The Windows Unified Data Storage Server console facilitates common management tasks for the various file services and the iSCSI target function. This console integrates several Microsoft Management Console (MMC) snap-ins to provide a single, unified management tool for the PowerVault NX1950 solution. Administrators can use a Web-based or remote desktop interface for remote management. Remote administration support also includes a Windows Server 2003 Remote Administration applet that enables browser-based management from non-Windows clients.

To provide a consistent interface for storage subsystem management and to avoid the need for device-specific management applications, the PowerVault NX1950 includes Microsoft Virtual Disk Service (VDS), which enables common configuration and management tasks such as provisioning storage resources like LUNs and disk volumes. The Share and Storage Management snap-in enables provisioning of the Dell PowerVault SAS or Dell/EMC Fibre Channel SAN storage supported by the PowerVault NX1950 solution. To enable this functionality, the PowerVault NX1950 is preconfigured with VDS providers that translate VDS requests into device-specific commands understood by the supported storage arrays. It also includes support for the Microsoft iSCSI Software Target VSS provider, which enables backup application servers to obtain application-consistent snapshots of data disks hosted by iSCSI targets on the PowerVault NX1950.

Building a high-availability Dell PowerVault NX1950 cluster

Clustering Dell PowerVault NX1950 systems using Microsoft Cluster Service—available only with Windows Unified Data Storage Server 2003 Enterprise Edition—helps enterprises create highly available services. The PowerVault NX1950 can be configured with Windows Unified Data Storage Server 2003 Enterprise x64 Edition, which provides support for clustering up to eight server nodes. Because the cluster hosts data on shared storage accessible to all cluster nodes, if a node fails, applications or services from that node fail over to another node and thereby help minimize service downtime.

Figure 1 shows a typical PowerVault NX1950 clustering configuration, which includes two or more Dell PowerVault NX1950 systems attached to the same shared back-end storage device. Administrators can configure high availability for services including CIFS, NFS, DFS namespaces, iSCSI targets, and print spoolers. Dell supports clustering with a specific set of server and storage hardware components; for a list of these components, visit www.dell.com/ha.

Choosing back-end cluster storage

When choosing back-end cluster storage, enterprises should take into account cost, scalability, and other factors. Dell PowerVault NX1950 clusters support both direct attach storage, such as the Dell PowerVault MD3000 enclosure for the integrated approach, and Fibre Channel–based SANs using Dell/EMC storage platforms for the SAN gateway.

The PowerVault MD3000 is a direct attach SAS storage enclosure that can support up to two PowerVault NX1950 head units as part of a cluster and help provide a cost-effective clustering option for file sharing. Storage capacity scalability is limited by the number of hard drives supported by the storage enclosure.

Dell/EMC Fibre Channel storage platforms enable a range of entry- to enterprise-level SAN implementations. The PowerVault NX1950 is designed to seamlessly integrate with existing supported SAN infrastructures and help enable storage capacity scaling. Scalability can be affected by the use of multiple PowerVault NX1950 clusters in a single SAN or multiple back-end storage devices for a single cluster. PowerVault NX1950 clusters can also serve as gateways to SAN-based storage, providing highly available file shares and highly available iSCSI targets with a single simplified management interface.

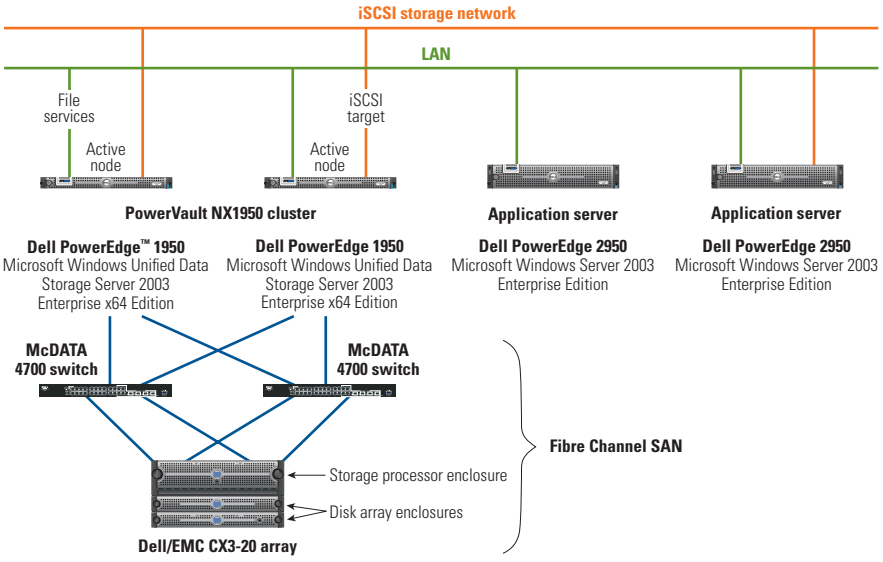


Figure 1. Typical high-availability Dell PowerVault NX1950 cluster

Configuring the cluster

When the Dell PowerVault NX1950 is first powered up, the Dell Initialization Wizard helps facilitate back-end storage initialization. After this configuration is complete, the PowerVault NX1950 Configuration Tasks window opens, which helps streamline basic configuration steps (see Figure 2). In addition to configuring Microsoft Services for NFS, administrators can use the tasks in the "Customize this system"

DELL IS CHANGING THE ECONOMICS OF STORAGE.

DELL POWERVAULT NX1950

THE NEXUS OF NETWORKED STORAGE, WHERE SIMPLICITY, SCALABILITY AND INTEROPERABILITY CONVERGE.

Dell™ PowerVault™ NX1950: A unified network storage solution, delivering advanced features for an "off the shelf" price

Flexible and versatile deployment: Combining the simplicity of NAS with the scalability of SAN.

- Supports NFS and CIFS protocols
- Manage both block and file data with a cost-effective iSCSI solution

Engineered for High Availability: Protect your critical data with a PowerVault NX1950 active/active cluster configuration, maintaining high system and data availability.

Advanced software capabilities included – no penalty for growth: PowerVault NX1950 comes bundled with point-in-time snapshots, remote replication, set quotas and multi-node clustering configurations, just to name a few. Grow your capacity without having to worry about increased licensing fees.

Simple unified management: Manage system and storage components through one easy-to-use management tool, Microsoft® Windows® Unified Data Storage Server 2003.

Microsoft® Windows® Unified Data Storage Server 2003

www.dell.com/NX

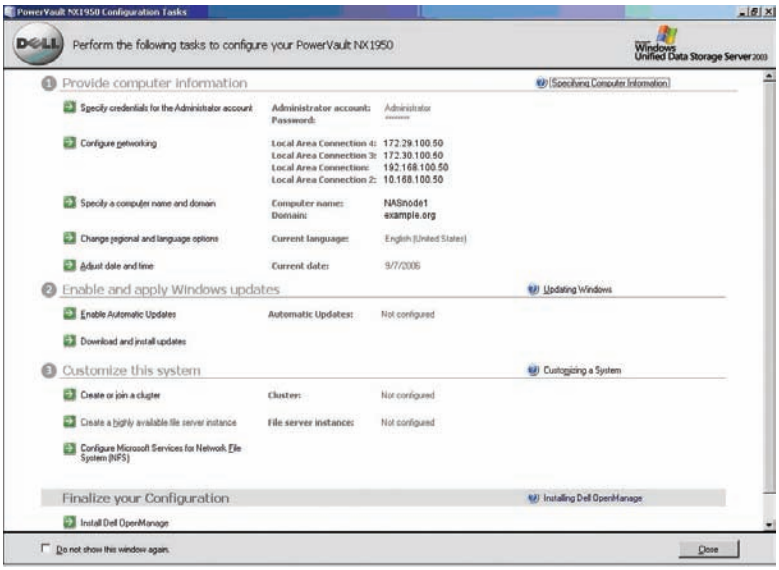


Figure 2. Dell PowerVault NX1950 Configuration Tasks window

section of this window to create or join a cluster and create a highly available file server instance.

Create or join a cluster. This task helps simplify the process of building a cluster. When run, the task presents a list of prerequisites; to help ensure that these prerequisites have been met, administrators should complete the initial configuration tasks in the order presented.

When this task is initiated from the first node in a cluster, it performs the following steps:

1. Creates a suitable LUN that resides on the shared back-end storage device and serves as the quorum resource for the cluster
2. Assigns this LUN to the first node in the cluster
3. Runs the Cluster Configuration Wizard, from which administrators can configure the cluster name and management IP address
4. Provides a dialog box that allows administrators to specify additional nodes, and assigns the quorum disk LUN to those additional nodes

LUN creation and assignment are performed using VDS, so administrators do not need to manually perform these operations on the back-end storage device or use the traditional management interfaces for provisioning.

When this task is initiated from additional nodes, the Cluster Configuration Wizard helps join the node to the existing cluster. After the task has been run on all nodes, the basic framework for a high-availability cluster is in place. The next step is to extend this framework and make it usable for CIFS and NFS file shares.

Create a highly available file server instance. This task automates the creation of cluster resource groups, which allow file shares to benefit from the fault tolerance of a high-availability networked storage cluster. Administrators specify a network name, IP address, and subnet mask for the highly available file server instance. They can then use the Provision Shares task in the Windows Unified Data Storage Server console to assign LUNs and configure highly available file shares within these cluster resource groups.

A highly available file server instance can run on only one node in a cluster. Regardless of the node on which the instance is running, any shares configured for the instance are available to clients through the specified network name and IP address. To help distribute the load generated by several file shares across the nodes, administrators can create a highly available file server instance on each node, then select different instances when creating highly available file shares from the Provision Shares task.

Implementing the cluster

The principal goal of a high-availability networked storage cluster is to reduce the possibility that a single point of failure can cause the data served by the cluster to become inaccessible to client systems. Microsoft Cluster Service allows the file- and block-level storage provided by the Dell PowerVault NX1950 to be highly available; if one node fails, the remaining nodes in the cluster assume control of the data and services that were running on the failed node, and the cluster continues to provide a uniform point of access for clients. The PowerVault NX1950 provides fault tolerance for the key networked storage features listed in Figure 3.

High-availability PowerVault NX1950 networked storage clusters can provide virtual servers configured as cluster resource groups. These virtual servers allow clients to access file shares or iSCSI targets in a consistent manner regardless of the node on which the services are physically running. Administrators can configure this single point of access for CIFS and NFS file shares, stand-alone DFS namespaces, and the iSCSI target service. In addition, quotas or file-screening policies that are configured on a file share are preserved regardless of which node hosts the share.

Deploying high-availability Dell PowerVault NX1950 networked storage clusters

The Dell PowerVault NX1950 storage server can provide highly available networked file services and includes support for common file-based sharing protocols and block-level storage through an iSCSI target service. Customized management consoles and remote administration capabilities help simplify storage provisioning and

Networked storage feature	High-availability deployment
CIFS and NFS file shares	Share assigned (using the Provision Shares task) to a highly available file server instance configured on the networked storage cluster
Quotas and file screening	
DFS namespaces	• Stand-alone DFS namespace configured as a resource on the networked storage cluster • Highly available file share published (using the Provision Shares task) to a DFS namespace
iSCSI target (block-level storage)	Microsoft iSCSI target service configured as a resource on the networked storage cluster
Print spooler	Print spooler resource configured on the networked storage cluster
Remote management	Microsoft Windows Unified Data Storage Server console (which provides provisioning capabilities and a link to Cluster Administrator) installed on each node; console is accessible from a Web browser using the Remote Desktop Protocol

Figure 3. Networked storage features and their high-availability deployments

monitoring tasks, and unified management features enable administrators to configure various back-end storage devices transparently. Deploying the PowerVault NX1950 can enhance the flexibility and manageability of enterprise networked storage infrastructures.

Kevin Guinn is a systems engineer in the High-Availability Cluster Development Group at Dell. His recent efforts have focused on business continuity and core network services. Kevin is a Microsoft Certified Systems Engineer and has a B.S. in Mechanical Engineering from the University of Texas at Austin.

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, databases, application performance, and cluster management. Ananda has a master's degree in Computer Science from Texas A&M University.

FOR MORE INFORMATION

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

Microsoft Windows Storage Server 2003:
www.microsoft.com/windowsserversystem/wss2003

Microsoft Windows Server 2003 file and storage services:
www.microsoft.com/windowsserversystem/storage/storservices.mspx

Microsoft Simple SAN initiative:
www.microsoft.com/windowsserversystem/storage/simplesan.mspx

Reprinted from *Dell Power Solutions,* February 2007.

Enabling Real-Time Protection for Microsoft Exchange

with Symantec Backup Exec 11d and Continuous Protection Server

Symantec® Backup Exec™ 11d for Windows Servers software, along with the Continuous Protection Server feature and Backup Exec Agent for Microsoft Exchange Server, can provide fast, flexible technology to help protect vital enterprise messaging data in Microsoft® Exchange environments. This article examines how this software can protect Exchange data on Dell™ PowerEdge™ servers with Dell PowerVault™ storage.

BY CHARLES BUTLER, KYON HOLMAN, AND CAROLINA MARTINEZ

Related Categories:

Backup, recovery, and archiving (BURA)

Dell PowerEdge servers

Dell PowerVault storage

Microsoft Exchange

Storage software

Symantec

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

E-mail has become an integral part of daily enterprise operations, and e-mail data requires effective, flexible backup and recovery. Unfortunately, the backup window available for e-mail systems—the limited period during which backup operations run—often shrinks as the demands placed on protection and recoverability grow. Although traditional tape-based backup is a critical component of a business continuance strategy, it is typically slower and can be less reliable than disk-based backup, including continuous disk-based data protection.

Using Symantec Backup Exec 11d for Windows Servers in conjunction with the Continuous Protection Server (CPS) feature and Backup Exec Agent for Microsoft Exchange Server enables real-time disk-based data protection while enhancing reliability, providing fast backup and recovery, and helping eliminate problems with backup windows. The Backup Exec Agent for Microsoft Exchange Server uses Backup Exec and CPS to provide continuous data protection and full, incremental, or differential backup and restore of embedded objects, attributes, and Microsoft Outlook® components. This tight integration enables administrators to recover e-mail messages or mailboxes from continuous or full backup jobs, helping greatly reduce recovery time for these components. More important, this solution provides administrators with the proper tools to help meet the real-time continuance requirements of their critical Exchange environments.

Understanding Continuous Protection Server

In typical enterprise data centers, backups are managed by IT administrators, who use a backup application to identify and restore files; this process is inaccessible to end users. Continuous data protection, however, can provide power and flexibility to both administrators and end users—enhancing the processes for protecting critical enterprise information.

CPS protects files continuously and in real time—whenever a file is created or changed, it is protected immediately. By using this feature, administrators can eliminate the full, incremental, and differential backups inherent in

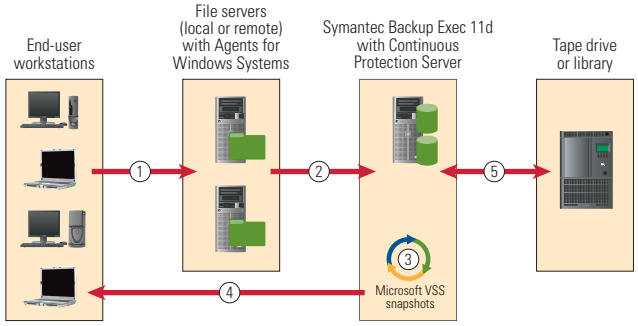


Figure 1. Backing up and restoring Windows-based file servers using Symantec Backup Exec 11d with Continuous Protection Server

tape-based backups. CPS runs on an existing Backup Exec media server or a separate server to allow maximum flexibility, and works by installing an Agent for Windows Systems (AWS) on one or more file servers and directing them to the CPS server. This agent is included in the Backup Exec Agent for Microsoft Exchange Server, and continuously tracks and records changes to files as they are created or updated; when a file is created or updated, the AWS transmits the changed data in the file to the CPS server. Pairing the CPS server with the internal disk storage on a server (such as a Dell PowerEdge 2900) or an external storage device (such as a Dell PowerVault MD3000) helps protect files continuously and eliminate backup windows.

Figure 1 illustrates how CPS functions within a Backup Exec environment:

1. End users save files to file servers.
2. The AWS on each server streams file changes to the CPS server.
3. Microsoft Volume Shadow Copy Service (VSS) snapshots provide point-in-time file recovery, and the CPS server offers smart management of these snapshots, allowing maximum flexibility in data protection.
4. End users retrieve previous versions of files using simple Web-based recovery.
5. The Backup Exec media server provides tape backups for long-term retention or disaster recovery.

Using Continuous Protection Server with Microsoft Exchange

The challenge of protecting Microsoft Exchange servers lies not in backing up the information stored on the servers, but in recovering e-mail messages, folders, and mailboxes if a failure occurs. Backup Exec with CPS provides a powerful way to meet this challenge, helping enterprises reliably protect and restore Exchange servers over networks.

As illustrated in Figure 2, an Exchange environment with Backup Exec and CPS consists of the following components:

- Exchange servers with the Backup Exec Agent for Microsoft Exchange Server (which includes the AWS)
- Symantec Backup Exec media server with CPS
- Tape and disk storage for Backup Exec backups
- Disk storage for CPS backups

The Backup Exec Agent for Microsoft Exchange Server is a key component of such an environment, providing integrated backups of Exchange databases with network backups to a Backup Exec media server. The Backup Exec Agent for Microsoft Exchange Server 2000 and Exchange Server 2003 provides the following features:

- **Granular restore:** The agent uses a granular restore feature for e-mail messages and folders backed up with the Exchange Information Store, providing enhanced performance for mailbox backup and restore. This feature enables administrators to restore individual e-mail messages and folders from full, copy, and incremental backups of the Information Store.
- **Continuous backup and comprehensive recovery:** Using the agent with CPS enables real-time backups and provides comprehensive recovery to any point in time for the Exchange Information Store, including the latest complete transaction log.
- **Recovery points:** Using the agent with CPS enables recovery points to run at intervals between the full database backups, so administrators can restore individual e-mail messages or folders to a point in time when the recovery point was created.
- **Selection of individual storage groups and databases:** The agent enables administrators to select individual storage groups, or individual databases within a storage group, for backup and restore jobs. They can also recover storage groups or databases by using the Recovery Storage Group feature of Exchange Server 2003 for non-snapshot backups.

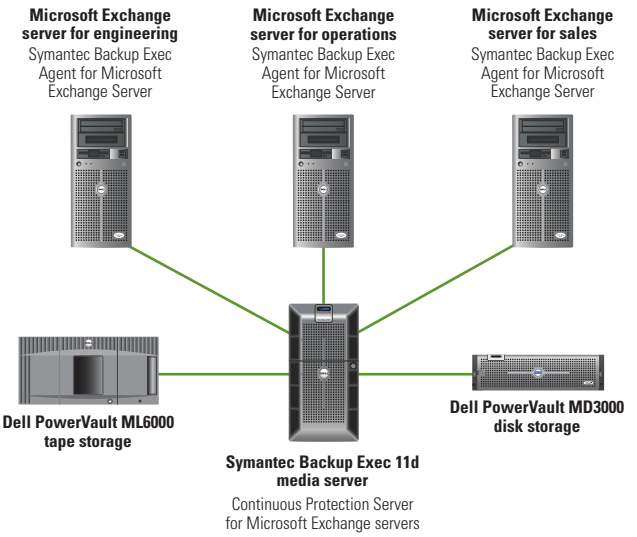


Figure 2. Implementing Symantec Backup Exec 11d and Continuous Protection Server in an example Microsoft Exchange environment

CPS integrates with Backup Exec and the Backup Exec Agent for Microsoft Exchange Server to provide continuous Exchange data protection through a CPS broker service installed with each Backup Exec Agent for Microsoft Exchange. This broker service allows administrators and end users to create Exchange log protection jobs to protect the Exchange transaction logs through the Backup Exec Administration Console. As a result, administrators and end users have a single interface to both CPS (for immediate data protection) and tape backup retention (for long-term storage).

Once the broker service is installed, administrators can create an Exchange log protection job to protect the Exchange transaction logs. These logs are part of the Backup Exec full Exchange backup set. Although Exchange protection jobs can only be created from the Backup Exec Administration Console, administrators can monitor the status of these jobs in the job log and report details from the Monitor view of the CPS Administration Console.

Configuring Continuous Protection Server and Microsoft Exchange

The Backup Exec 11d media server supports the following operating systems:

- Microsoft Windows® 2000 Server with at least Service Pack 4 (SP4) and Update Rollup 1 for SP4
- Microsoft Windows Server® 2003 with at least SP1
- Microsoft Windows Server 2003 Release 2 (R2)
- Microsoft Windows XP Professional with at least SP2

To support continuous protection, the following components are required on the Microsoft Exchange server:

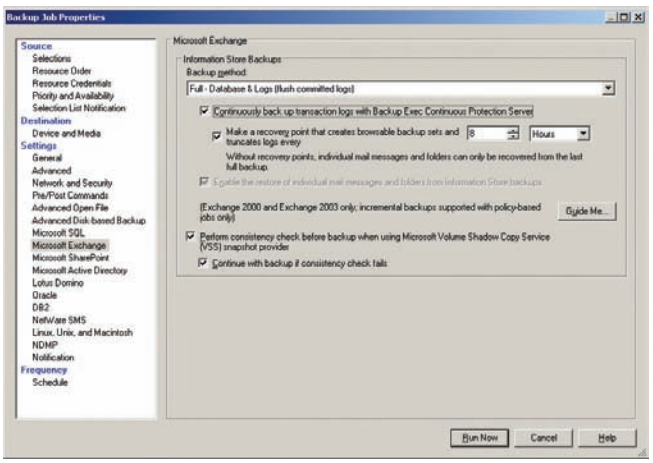


Figure 3. Configuring backup settings for Microsoft Exchange through the Backup Job Properties dialog box in Symantec Backup Exec 11d

- Windows 2000 Server or Windows Server 2003
- Exchange Server 2000 with SP3, Exchange Server 2003, or Exchange Server 2007
- Backup Exec Agent for Microsoft Exchange Server (which includes the AWS)[1]

The Backup Exec Agent for Microsoft Exchange Server must be licensed and installed on the Backup Exec media server along with the CPS Administration Console. Both agents must use the Backup Exec service account, and this account must have domain and local administrator rights and be able to query the location of the transaction logs on the local active Exchange server. In addition, the CPS media server (the server running Backup Exec with CPS) and the CPS enterprise server (the Exchange server being protected by CPS) must be in the same domain, unless the domains are trusted.

Creating continuous Microsoft Exchange protection

Creating continuous Microsoft Exchange protection requires regular full Exchange backups. Administrators enable continuous backups by selecting the "Continuously back up transaction logs with Backup Exec Continuous Protection Server" check box in the Backup Job Properties dialog box for Microsoft Exchange settings in Symantec Backup Exec 11d (see Figure 3). Backup Exec then protects the Exchange Information Store with a recurring full backup to a backup-to-disk folder on a local NT File System (NTFS) volume. Backup Exec also continuously replicates the Exchange transaction logs to a backup-to-disk folder, providing real-time protection.

As part of its continuous protection, Backup Exec can make recovery points at intervals specified by administrators. Recovery points create backup sets that can be browsed from the Restore view in Backup Exec and used to recover e-mail messages or folders from a point in time when either a full backup was run or a recovery point was made. To help control log growth, Backup Exec also truncates the transaction logs each time a recovery point is made.

Restoring Microsoft Exchange objects from Continuous Protection Server backups

The full backups and replicated transaction logs created by Backup Exec enable administrators to comprehensively restore data from the Microsoft Exchange Information Store to any point in time, including the latest complete transaction log. They can also restore the entire Information Store from a full backup or recovery point, the latest complete transaction log, or any point in time between full backups or recovery points (see Figure 4).

To restore the entire Information Store to the time of a full backup or recovery point, administrators can do the following:

[1] For more information and a list of the latest supported versions of Microsoft Windows and Microsoft Exchange Server, visit www.backupexec.com or see the Backup Exec 11d and CPS software compatibility lists at seer.support.veritas.com/docs/284493.htm and seer.support.veritas.com/docs/285788.htm.

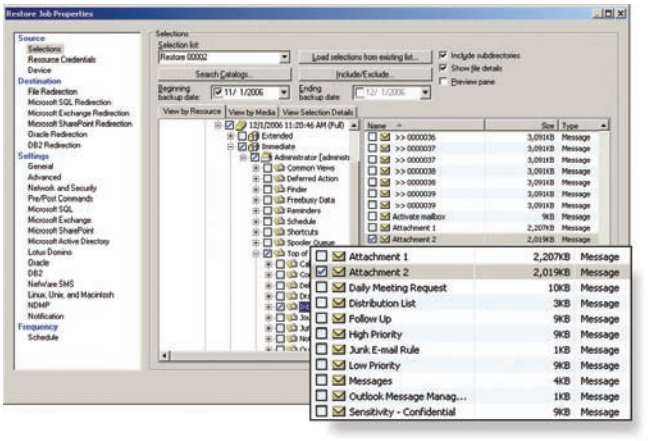


Figure 4. Restoring individual messages using the Restore Job Properties dialog box in Symantec Backup Exec 11d

1. Select the backup sets from the full backup or recovery point that contains the desired point in time.
2. In the Restore Job Properties dialog box for Microsoft Exchange settings, click "Purge existing data and restore only the databases and transaction logs from the backup sets."

To restore up to the latest full transaction log, administrators can do the following:

1. Select the backup sets from the last full backup or recovery point.
2. In the Restore Job Properties dialog box for Microsoft Exchange settings, click "Restore all transaction logs; do not delete existing transaction logs (no loss restore)."

To restore to a point in time between full backups or recovery points, administrators can do the following:

1. Select the backup sets from any full backup or recovery point and specify the point in time.
2. In the Restore Job Properties dialog box for Microsoft Exchange settings, click "Restore all transaction logs until point in time; skip transaction logs after this time."
3. Specify the point in time.

Following best practices for continuous Microsoft Exchange protection

When using continuous protection as part of a Microsoft Exchange backup strategy, administrators should follow these best practices:

- Symantec recommends backing up only one Exchange server for each continuous backup job. Administrators should create a separate selection list for each Exchange server resource.
- When copying backup sets to tape for off-site storage, administrators should create a job to duplicate backup sets. They can configure the job to copy the backup sets to tape after each occurrence of the full backup job.
- If the job template is in a policy, administrators should create a duplicate backup-sets template to copy the backup sets after the last incremental backup and before the full backup to tape. If necessary, they can also create a copy job to run before the full backup to copy all transaction logs as well as the full backup sets to tape.
- Circular logging in Exchange should be disabled.
- The "Enable the restore of individual mail messages and folders from Information Store backups" option is automatically selected when using the "Continuously back up transaction logs with Backup Exec Continuous Protection Server" option. These settings enable administrators to restore individual e-mail messages or folders from a CPS backup.

Protecting critical enterprise messaging data

Symantec Backup Exec 11d for Windows Servers and Continuous Protection Server enable administrators to implement real-time disk-based data protection for Microsoft Exchange servers, and allows both administrators and end users to flexibly restore files from backups and restore points. Using these components as part of a comprehensive enterprise backup strategy helps protect critical Exchange messaging data.

Charles Butler is a technical product manager in the Data and Systems Management Group at Symantec. He has a B.S. in Electrical and Computer Engineering from the University of Colorado at Boulder and an M.B.A. from St. Edward's University.

Kyon Holman is a lead software engineer on the Tape Storage team in the Dell Enterprise Product Group. He has a B.S. in Computer Science from the University of Michigan at Ann Arbor and an M.S. in Software Engineering from the University of Texas at Austin, and is currently pursuing an Executive M.B.A. from the University of Texas at Austin.

Carolina Martinez is a storage product test engineer in the Dell Enterprise Product Test Group. She has a B.E. in Computer Engineering and Applied Mathematics from the State University of New York at Stony Brook.

FOR MORE INFORMATION

Symantec Backup Exec:
www.backupexec.com

Dell storage:
www.dell.com/storage

Tiered Storage Consolidation

Using Dell/EMC CX3 UltraScale Series Arrays and EMC Software

Tiered storage consolidation can provide an efficient, cost-effective way to manage growing data volumes while helping simplify data management, increase storage utilization and availability, and enhance backup, recovery, and archiving processes. This article discusses implementing tiered consolidation using Dell/EMC CX3 UltraScale™ series storage arrays and EMC® Navisphere® Quality of Service Manager software.

BY KATIE CURTIN-MESTRE

Related Categories:

Dell/EMC storage

EMC

Fibre Channel

Storage area network (SAN)

Storage consolidation

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

Enterprises of all sizes struggle with the challenges of storing critical data in ways that not only protect it, but also ensure its availability to users and applications when and where they need it. In a survey of IT managers conducted by EMC Corporation, respondents cited backup, business continuity and disaster recovery, security, and long-term archiving as top storage concerns (see Figure 1). To properly address these concerns, enterprises must manage large volumes of data and multiple applications, scaling up their storage environments to meet increasingly complex capacity, performance, and availability requirements. Such an approach can lead to a heterogeneous collection of storage subsystems from different vendors added as needed over an extended period of time—which, taking into account both storage resources and the necessary dedicated staff, is typically not only time-consuming to manage and difficult to scale, but expensive as well.

Storage consolidation provides a simple, cost-effective way to meet growing requirements for data sharing, high performance, and high availability by centralizing storage resources among numerous application servers. Storage consolidation makes up a significant part of general IT consolidation: IDC predicts that by 2009, more than 18 percent of a US$24.7 billion worldwide IT consolidation market will represent spending on storage consolidation.[1] The potential benefits of storage consolidation include enhanced storage utilization, increased availability, centralized management, and reduced total cost of ownership, in addition to streamlined data

[1] "Worldwide IT Consolidation Market Opportunity," by IDC, Doc #201996, June 2006.

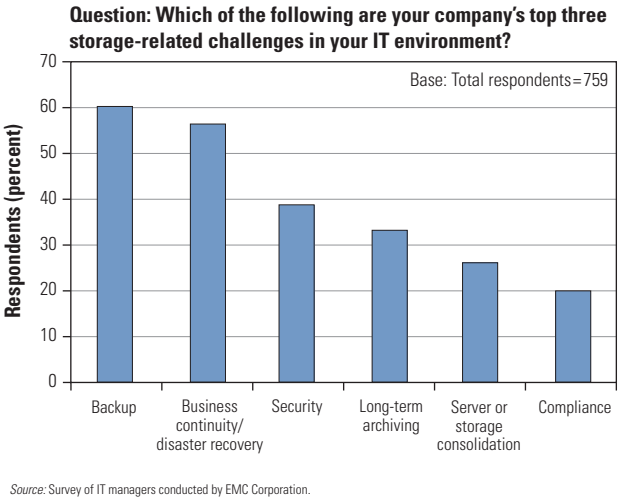


Figure 1. Survey responses on top storage-related challenges

protection processes through reduced backup and recovery windows and enhanced disaster recovery.

Although consolidation can improve the way enterprises store, protect, and access data, treating all data equally can result in suboptimal performance: as more and more applications reside on the same array, associating the most widely used information with the best-performing hardware can become difficult, preventing enterprises from taking advantage of the potential benefits of storage consolidation. *Tiered storage* provides a way to implement storage consolidation that helps avoid these problems.

Consolidating with tiered storage and quality-of-service tools

Shrinking budgets and increasing data volumes can combine to pressure IT organizations to maximize storage resources in an efficient, cost-effective environment. Implementing tiered storage means maintaining storage devices of varying performance and cost characteristics within the same or multiple storage systems, and typically represents the first step toward comprehensive storage consolidation. This setup gives administrators the flexibility to efficiently utilize resources by aligning hardware levels with information value—for example, by reserving high-performance disk drives and advanced software functionality for critical applications to help maximize the performance and availability of those applications.

If properly implemented, a tiered architecture can not only improve service levels for critical data, but also reduce overall costs by avoiding the use of the most expensive and highest-performance systems for noncritical data. Different categories of data are assigned to different types of storage media, with categories based on criteria such as level of protection needed, performance requirements, and frequency of use. In a three-tier system, for example, tier 1 data might include critical, recently accessed, or highly sensitive files and be stored on the most expensive, highest-quality, fastest-performing media in the environment. Tier 2 data might include financial, seldom-used, or moderately sensitive files and be stored on media that is both less expensive and lower-performing than the tier 1 media, while tier 3 data might include event-driven, rarely used, or nonsensitive files and be stored on media that is less expensive and lower-performing than the tier 2 media.

Understanding Fibre Channel and iSCSI storage area networks

Storage area networks (SANs) provide a widespread method for storing and managing enterprise data. Advantages of SANs can include high performance, the ability to interconnect storage devices with associated servers, and support for disk mirroring, backup and restore, data archiving and retrieval, data migration between storage devices, and other advanced storage functionality.

Two widely used standards for storage connection in a SAN environment are Fibre Channel (FC) and Internet SCSI (iSCSI). The FC standard, designed for connecting servers to shared storage systems, represents a mature storage technology that can provide high performance and availability and offers both widespread existing deployment and massive growth potential. The maximum host throughput and bandwidth is typically much higher for hosts connected through FC than for those connected through iSCSI.

If properly implemented, a tiered architecture can not only improve service levels for critical data, but also reduce overall costs by avoiding the use of the most expensive and highest-performance systems for noncritical data.

The iSCSI standard—an IP-based standard for linking data storage devices over a network and transferring data by carrying SCSI commands—is extremely flexible, enabling data transmission over LANs, wide area networks (WANs), or the Internet as well as location-independent data storage and retrieval. It typically costs less than FC: according to the Sageza Group, iSCSI typically costs between 18 percent and 40 percent less than a comparable FC configuration, depending on the hardware and number of connected servers.[2]

For midsize enterprises, FC is typically appropriate for decision support systems, backup-to-disk operations, and streaming

[2] "EMC Energizes iSCSI Deployment for Windows Servers," by the Sageza Group, February 14, 2005, www.emc.com/analyst/pdf/SN2-14-05_iSCSI.pdf.

media applications, while iSCSI is typically appropriate for e-mail and online transaction processing (OLTP) applications. For large enterprises with demanding requirements, FC can offer the performance necessary for key applications, while iSCSI can complement FC deployments and help cost-effectively consolidate applications not using FC. Organizations of all sizes are becoming aware of the benefits that iSCSI can deliver at a much lower cost than FC: a survey of IT professionals conducted by EMC Corporation revealed that 11 percent of midsize enterprises had deployed iSCSI and 50 percent were evaluating it or planned to evaluate or deploy it in the next 18 months, while 21 percent of large enterprises had deployed iSCSI and 59 percent were evaluating it or planned to evaluate or deploy it in the next 18 months.[3]

Adding quality-of-service tools

Although tiered storage consolidation can be an effective cost-savings method, a consolidated environment can be complex—and implementing and managing it while meeting numerous performance goals in a multi-application environment can be both labor-intensive and expensive without the necessary software functionality. Enterprises typically must either add dedicated staff or rededicate staff from critical tasks, and because of these and other cost-related drawbacks, some enterprises may overprovision storage resources, preventing them from taking advantage of the potential savings benefits of storage consolidation. Fortunately, some storage consolidation implementations can simplify and automate quality-of-service (QoS) management, helping protect the enterprise value of consolidated storage.

QoS tools enable administrators to easily, reliably, and efficiently deliver high service levels for critical applications by automating the dynamic process of allocating system resources. Administrators can monitor performance, then set service-level requirements and performance targets for specific applications within an array, effectively increasing the information's enterprise value and helping ensure appropriate service levels. They can also set limits on lower-priority applications and schedule policies to run at different intervals. Different tools allow different levels of control; more granular control over classifying I/O requests typically increases administrators' ability to optimize application performance.

An effective quality-of-service tool can isolate the root cause of the problem and help determine whether the storage system was actually the cause, allowing administrators to accurately and effectively address problems as they occur and resolve them quickly.

QoS tools can not only help improve performance, but also enable administrators to determine the cause of performance problems when they occur. In a complex storage environment, application performance problems can be reflexively blamed on storage with little or no evidence. An effective QoS tool can isolate the root cause of the problem and help determine whether the storage system was actually the cause, allowing administrators to accurately and effectively address problems as they occur and resolve them quickly.

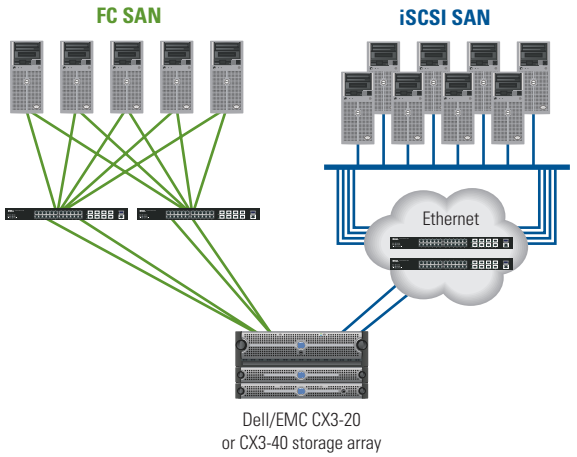


Figure 2. Storage consolidation with a Dell/EMC CX3 series storage array providing Fibre Channel and iSCSI connectivity

Implementing flexible tiered storage with the Dell/EMC CX3 series and EMC software

FC and iSCSI each have their own advantages and disadvantages, and neither can meet all the requirements of every storage environment on its own. In the past, enterprises planning to consolidate storage have had to choose between the high-capacity performance of FC and the cost-effectiveness and ease of use of iSCSI, and those wanting to add iSCSI capabilities needed to add a separate system to their environments. Unfortunately, the costs of doing so typically outweighed the benefits, leading some enterprises to avoid consolidating lower-tier applications altogether and leaving many applications on server-resident storage.

The Dell/EMC CX3 UltraScale series of storage arrays, however, supports both standards: all CX3 models support FC, and the CX3-20 and CX3-40 support both FC and iSCSI. Deploying these storage arrays enables enterprises to create efficient, cost-effective storage environments; take advantage of the potential benefits of tiered

[3] Survey of IT professionals at 1,318 enterprises worldwide conducted by EMC Corporation from January to March 2006.

Servers

Storage

Systems Management

Services

SQL Server 2005

How Dell Does IT

No one understands your need for database performance and availability better than Dell™ and Microsoft®. Microsoft SQL Server™ running on Dell PowerEdge™ servers are part of the foundation for Dell's retail website, providing the brains behind a number of our key customer-facing applications. This powerful combination helped to improve our application management, disaster recovery and server management, and is expected to increase application performance 1.4-times over the previous platform.

To help you experience the greatest potential of your database environment, Dell offers a complete SQL Server solution including servers, storage, systems management, services, and the software itself.

Visit www.dell.com/sqlmag for the complete story on how Dell IT uses SQL Server 2005.

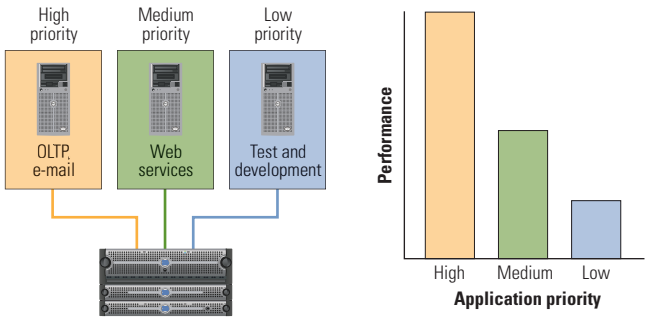


Figure 3. Different priority levels set for various applications using EMC Navisphere Quality of Service Manager

storage consolidation, including high performance, availability, and scalability; and meet a wide range of service-level requirements. The CX3-20 and CX3-40 provide the flexibility to deploy FC and iSCSI in one system, allowing enterprises to consolidate multiple application tiers into a single cost-effective storage system (see Figure 2). For example, enterprises could use the CX3-20 array's FC connectivity for decision-support and backup-to-disk applications, and the array's iSCSI connectivity to help consolidate storage for departmental servers running small OLTP applications and other information both inside and outside the data center. The scalability of the CX3-20 and CX3-40 arrays also provides flexibility for enterprises deploying FC that may also need to support iSCSI in the future.

The CX3 series includes the cost-effective CX3-20, the midrange CX3-40, and the most powerful of the CX3 series, the CX3-80. These arrays allow enterprises to consolidate just a few or hundreds of applications in a cost-effective way, and provide the flexibility to mix high-performance and cost-effective drives in a tiered storage environment to meet different service levels. The CX3 series' UltraScale architecture, based on PCI Express technology, enables high bandwidth and low latency, supports up to 4 GB/sec throughput, and recognizes and simultaneously supports devices of mixed speeds. Its virtual logical unit technology also enables dynamic migration of data between multiple storage tiers within a CX3 array.

Deploying EMC Navisphere Quality of Service Manager

EMC Navisphere Quality of Service Manager (NQM) software is a QoS tool designed to help enterprises maximize their investment in Dell/EMC CX3 arrays. It enables enterprises to prioritize applications and set specific performance targets, which can enhance productivity, help simplify management, and help administrators meet service-level requirements.

NQM includes the following features:

- **Performance monitoring:** NQM monitors individual application performance and provides logical views of both storage arrays and specific applications. These capabilities allow administrators to evaluate arrays to determine actual service-level requirements, and presents a clear picture of the performance service levels possible in a specific environment.
- **Performance targets:** NQM enables administrators to set targets for high-priority applications and limit resources for low-priority applications (see Figure 3). It includes a scheduler that allows them to adjust service-level requirements dynamically, giving priority to certain applications during specified periods—for example, giving high priority to an e-mail application during the day to improve response time during peak usage, then giving high priority to a backup application at night.
- **Diagnostics:** NQM provides advanced diagnostic capabilities to help evaluate application performance as it relates to storage arrays. If an application experiences performance problems, these features help administrators isolate the root cause and determine whether the problem is related to storage.

Creating efficient, cost-effective enterprise storage environments

Managing and storing increasing volumes of information presents many challenges, and tiered storage consolidation is a viable tool for overcoming those challenges. But to take full advantage of its potential benefits, enterprises must deploy systems with performance, price, and connectivity support appropriate to their storage environments.

The Dell/EMC CX3 series enables enterprises of any size to take advantage of the potential benefits of tiered storage consolidation in a cost-effective way, and provides flexible support for both Fibre Channel and iSCSI connectivity to allow enterprises to deploy each technology where it is most appropriate. The addition of EMC Navisphere Quality of Service Manager provides sophisticated management abilities designed to enhance productivity, simplify management, and help meet performance service-level requirements. In combination, the scalability, manageability, and cost-effectiveness of this hardware and software can provide an efficient platform for managing information throughout the enterprise.

Katie Curtin-Mestre is the director of product marketing for EMC CLARiiON® storage.

FOR MORE INFORMATION

Dell/EMC storage:
www.dell.com/emc

How Dell Accelerates Product Development Worldwide

Using PTC Pro/ENGINEER and Windchill Software

For more than a decade, Dell has designed and developed industry-leading hardware platforms using PTC® Pro/ENGINEER® Wildfire™ computer-aided design (CAD) software running on Dell Precision™ workstations. To advance its worldwide data management and collaboration, Dell recently deployed PTC Windchill® content and process management software on Dell™ PowerEdge™ servers, Dell/EMC storage area networks, and Dell PowerVault™ storage systems. This approach streamlines the global product development process—helping to improve business agility and accelerate time to market.

BY STEVE COLE, STEVE SHELTON, JERRY RODRIGUEZ, AND LARRY SALINAS

Related Categories:

Dell Precision workstations

Implementation study

Professional workstations

PTC

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

In today's world, the development of manufactured products is often a complex affair that traverses the globe. Not only are manufactured products themselves becoming increasingly sophisticated, but the companies that make such products increasingly rely on geographically distributed development organizations to help design them. To be successful in a global environment, however, a company must efficiently and effectively manage product development both within and across its design facilities. Failure to do so can lead to delayed product releases, missed market opportunities, and ultimately lost revenue.

Dell is no stranger to global product development. In fact, Dell has five globally distributed design centers, eight original design manufacturers (ODMs) in nearly a dozen locations, and numerous remote team members all contributing to the design and development of its industry-leading hardware platforms. Despite the complexity of this worldwide product development effort, Dell has a consistent track record of meeting its product release dates, and Dell products have been regularly hailed as technology leaders.

For more than a decade, Dell has relied on PTC Pro/ENGINEER Wildfire computer-aided design (CAD)

software running on Dell Precision workstations to design complex products and manage product information and development processes effectively. Recently, to advance communication and collaboration across its geographically distributed design centers, Dell deployed PTC Windchill content and process management software on Dell PowerEdge servers, Dell/EMC storage area networks (SANs), and Dell PowerVault storage systems. Windchill facilitates the central management of product data, streamlines and automates processes, and allows expanded yet secure access to product information—all of which contribute to increased efficiency and reduced time to market. In the future, Dell plans to extend its deployment of Windchill to facilitate comprehensive product life-cycle management (PLM) from concept all the way to servicing.

Windchill facilitates the central management of product data, streamlines and automates processes, and allows expanded yet secure access to product information—all of which contribute to increased efficiency and reduced time to market.

This article describes how Dell uses PTC products and solutions to enhance the product development process throughout its organization—including Pro/ENGINEER Wildfire CAD solutions for product design and Windchill content and process management software for collaboration. By running PTC software on its own standards-based workstations, servers, and storage systems, Dell is laying the foundation for a comprehensive, highly scalable PLM system.

Using PTC Pro/ENGINEER Wildfire for product design

Dell has used PTC Pro/ENGINEER software successfully in regional product design centers for more than a decade. Currently, Dell deploys hundreds of Dell Precision M90, Dell Precision 490, and Dell Precision 690 workstations running the Microsoft® Windows® XP OS and Pro/ENGINEER Wildfire CAD software to design and develop the entire Dell product line—including servers, workstations, desktops, notebooks, printers, and even televisions (see Figure 1). Pro/ENGINEER Wildfire offers several key features, including sophisticated 3-D modeling and simulation capabilities, an integrated architecture that supports the design process all the way from concept to manufacturing, and automated generation of associative tooling design, assembly instructions, and machine code.

By deploying Pro/ENGINEER Wildfire on standards-based Dell Precision workstations running Microsoft Windows, Dell can scale product design and development efforts quickly and cost-effectively in response to changing business requirements around the globe.

Targeting efficiency improvements

Rapid growth required the Dell regional design centers to work together as efficiently as possible to develop new products. However, the legacy system Dell had relied on to manage its product development efforts was not designed to support global collaboration. As a result, communication across design centers and with ODMs had become too time-consuming to keep up with business requirements to reduce both product development cycles and time to market.

In particular, because each design center managed its own data and processes, sharing product information and coordinating management tasks interactively could be challenging. For example, to share CAD data between different design centers, the legacy system relied on nightly FTP or e-mail file transfers from one design center to another—meaning that an FTP or e-mail failure could delay the transfer of critical data. Also, management tasks such as routing, review, and approval were performed on an ad hoc basis, which also had the

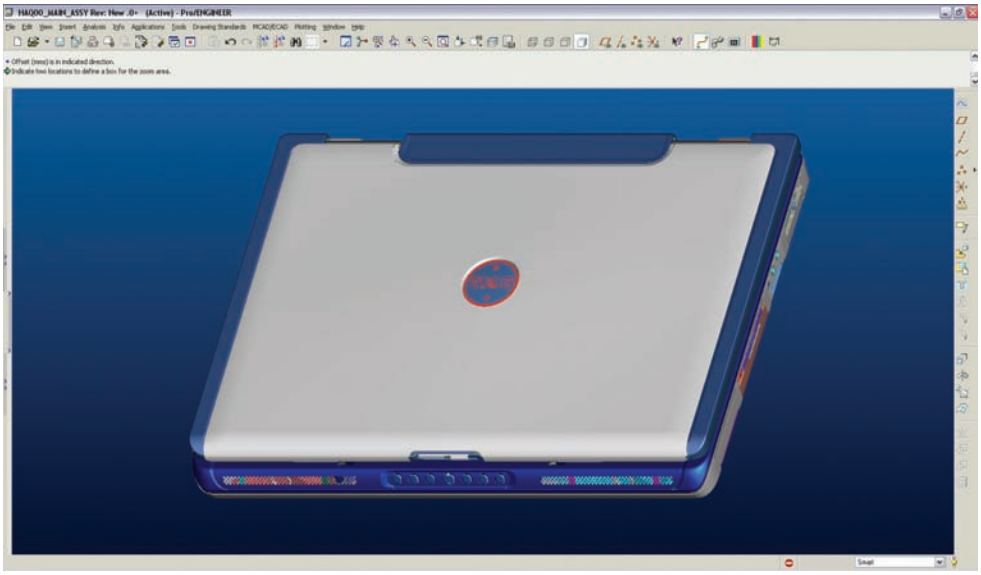


Figure 1. Dell design rendering of a Dell Precision M90 workstation chassis created using PTC Pro/ENGINEER Wildfire CAD software

potential to result in delays. And because the legacy system did not provide a standard way of managing global access to product information, accessing that information could be slow and cumbersome for partners and non-engineers.

Implementing PTC Windchill content and process management software

To help make its overall product design and development efforts as efficient as possible, Dell implemented PTC Windchill 8.0 content and process management software in August 2006. Windchill, a suite of products specifically designed to facilitate product development in a global environment, offers several compelling features and capabilities, including the following:

- **Single product-data repository:** Windchill uses a single repository for all product data, including both electronic and mechanical CAD (ECAD and MCAD) models, software builds, bills of material, configurable product structures, and other related documents—avoiding the need to manually transmit product data using FTP or e-mail from site to site.
- **Comprehensive process management:** Windchill facilitates process management throughout the product life cycle, allowing organizations to streamline key functions such as change management, configuration management, project management, and release management.
- **Sophisticated security and access controls:** Windchill includes comprehensive security and access controls, allowing expanded yet managed access for partners and non-engineers such as product marketing managers.
- **Global scalability:** Windchill is Web based, enabling engineers and non-engineers alike to access product data in a simple, easy-to-use way from anywhere in the world without having to install custom client software.
- **Out-of-the-box integration with PTC Pro/ENGINEER Wildfire:** Windchill software is fully integrated with Pro/ENGINEER Wildfire, helping eliminate the need for manual software integrations.

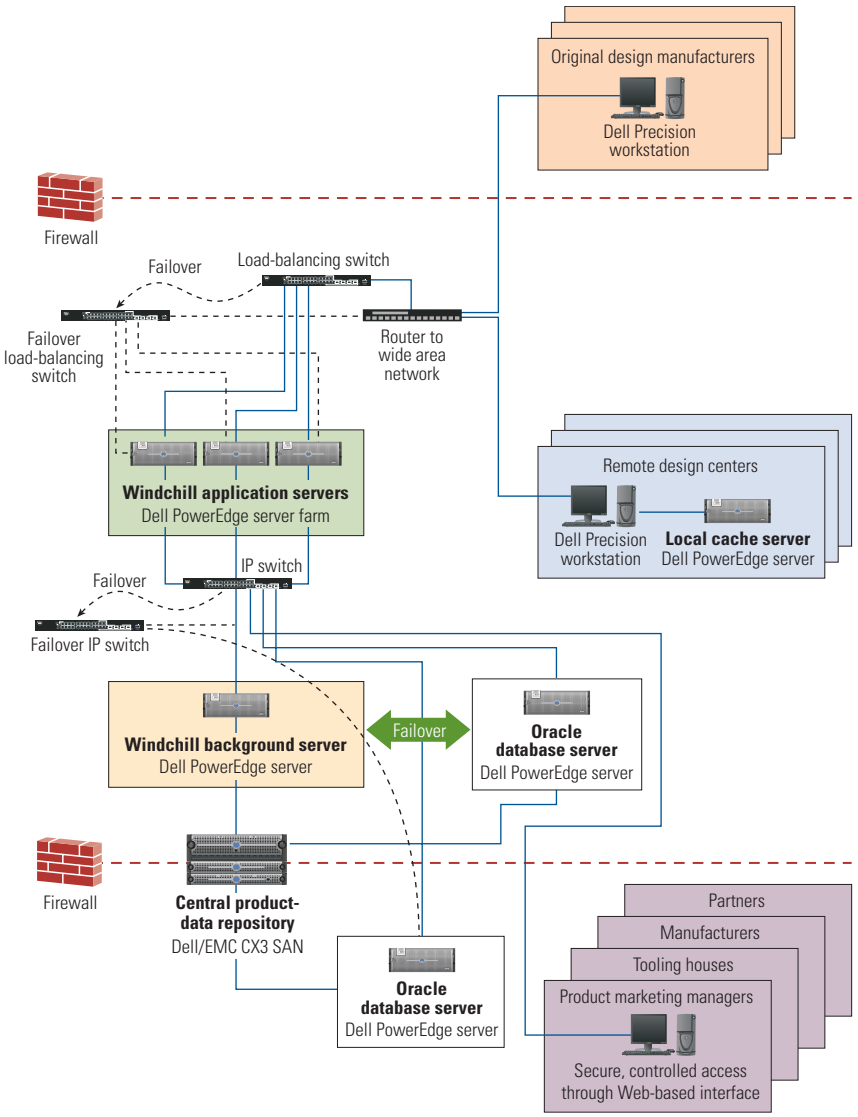


Figure 2. Example design center deployment to facilitate global collaboration

Dell deployed Windchill at its corporate design center in Austin, Texas; users at remote design centers, ODM facilities, and other locations can access the system through a simple Web-based interface. Although product data is centralized at the corporate design center in Austin, Dell utilizes PTC replication technology within Windchill to push large data sets to its design centers' cache servers to help increase regional performance.

Deploying PTC Windchill on Dell PowerEdge servers and a Dell/EMC SAN

Figure 2 shows an example design center deployment that would help facilitate the global collaboration necessary for a large product development organization such as Dell's. It comprises a farm of Dell PowerEdge servers running the Red Hat® Enterprise

Linux® OS and PTC Windchill PDMLink® software for product-data management and PTC Windchill ProjectLink™ software for project management. Supporting the Windchill applications are two additional PowerEdge database servers, each running the Red Hat Enterprise Linux OS and Oracle® database software.

To help maximize system uptime, key elements of the infrastructure are duplicated to facilitate failover in case of system failure. For example, the main Oracle database server is replicated to a backup Oracle database server, also deployed on a PowerEdge server. In addition, load-balancing switches not only optimize the performance of the Windchill application servers, but also help ensure application server availability. Finally, the load-balancing and IP switches themselves are duplicated.

A Dell/EMC CX3 SAN provides a central product-data repository to help manage the several terabytes of CAD data and other product information typically generated by a large product development organization. To support fast and efficient access to product data at remote design centers, product data is cached locally at each design center on PowerEdge servers.

Accelerating time to market

Since implementing PTC Windchill, the Dell product development team has reported improvements in efficiency that are expected to help shorten time to market for new products and reduce overall global development costs. Thanks to a centralized product-data repository and streamlined process management, Dell engineers can communicate and collaborate quickly and effectively with colleagues around the world using standardized, automated processes for routing, review, and approval.

Moreover, the advanced security and access features of Windchill enable Dell to extend access to product information beyond the engineering department, facilitating the review of new products earlier in the design cycle and enabling faster creation of prototypes than could be accomplished with the legacy system. Now, non-engineers such as product marketing managers can access product data from a simple, secure, Web-based interface—avoiding the need for engineers to spend time manually extracting information. And Dell partners such as manufacturers and tooling houses have controlled, Web-based access to the product information they need—enabling them to easily and rapidly manufacture product prototypes.

Deploying Windchill on high-performance, industry-standard Dell PowerEdge servers in a centralized location enables the Dell team to simplify operations, improve resource utilization, and scale in cost-effective increments—enhancing productivity while easing administration and maintenance tasks. In this way, Dell anticipates that Windchill will contribute to significant cost savings in the coming years. For example, Windchill uses the Web for client access to centralized resources, so Dell does not have to maintain separate systems at each regional design center and consequently does not have to install and maintain client applications at each regional design center. And because the Windchill architecture is highly scalable, Dell is confident that it will support aggressive growth of the Dell product development organization.

Enhancing business agility

Ultimately, Dell plans to expand its deployment of PTC Windchill to facilitate comprehensive PLM from concept all the way to servicing. With PLM, Dell hopes to integrate its entire range of off-the-shelf and custom-built product development applications into a single solution that will enable Dell and its partners to manage global product development more effectively than ever before.

A PLM solution based on PTC Windchill software and Dell hardware allows Dell to quickly and effectively incorporate customer feedback into new products; improve collaboration across departments such as engineering, customer service, and sales; and further streamline processes such as regulatory compliance and prototype creation—all of which Dell expects will lead to improved customer satisfaction and reduced time to market.

Steve Cole is the PTC sales technical account manager for Dell, working with Dell to develop technical requirements and solutions for new business. Steve is also closely involved in ongoing implementations as a technical adviser. He has a B.S.M.E. from the University of Texas at Austin.

Steve Shelton is the PTC global services program manager and is responsible for the successful implementation of PTC applications at Dell. Steve works with Dell management to coordinate and deploy multiple projects involving PTC software. He has many years of experience with successful enterprise program deployments.

Jerry Rodriguez is an IT strategist on the Dell Product Group's IT team, and is involved in Dell PLM initiatives. He is a graduate of Texas A&M University and has been in the high-tech industry for nearly 20 years.

Larry Salinas is the computer-aided engineering and IBM® Rational ClearCase manager for Dell, working with engineering teams to develop technical requirements and solutions for product design. He has a B.S.E.E. from the University of Texas at Austin.

FOR MORE INFORMATION

Dell Precision workstations:
www.dell.com/precision

PTC Pro/ENGINEER:
www.ptc.com/appserver/mkt/products/home.jsp?k=403

PTC Windchill:
www.ptc.com/windchill

Small Business

Dell recommends Windows® XP Professional

GEARBOX APPROACHED US WITH A PROBLEM. They were a game studio looking for a leg up on the competition. They needed an edge. And they got it, with a little help from Dell and Adobe. We gave them options like Dell Precision™ Workstations, high-performing systems equipped with Intel® Quad Core Processors that provide the power necessary for multi-threaded multimedia applications and customizable with advanced graphics, RAID hard drive support, and dual monitor capability. And through our partnership with Adobe, Gearbox could get the powerful Adobe OpenHD certified HDV solution featuring the comprehensive post-production tools of Adobe Production Studio. That meant access to the exceptional quality and resolution of HD at an affordable price. Which also meant that with their technology problems solved, Gearbox was free to churn out original blockbusters like 2005's Brother In Arms: Earned in Blood. Small business solutions designed for one company in mind. Yours.

DELL PRECISION™ M90
MOBILE WORKSTATION

***NEW* DELL PRECISION™ 490**
WORKSTATION

Adobe Production Studio Standard
Software, documentation or packaging may vary from retail version.

Adobe | OpenHD

Dell recommends Adobe® software with Dell Precision™ Workstations.

Adobe® Production Studio Standard
Adobe offers the essential post-production toolset.

- Package includes: Adobe After Effects® 7.0 Standard, Adobe Premiere® Pro 2.0 and Adobe Photoshop® CS2, Adobe Dynamic Link and Adobe Bridge.

Adobe® Production Studio Premium
Adobe offers a complete post-production solution.

- Package includes: Adobe After Effects® 7.0 Professional, Adobe Premiere® Pro 2.0, Adobe Photoshop® CS2, Adobe Audition® 2.0, Adobe Encore® DVD 2.0 and Adobe Illustrator® CS2 software, Adobe Dynamic Link and Adobe Bridge. **Call for Dell Pricing.**

Adobe

DELL Purely You

Call: M-F 7a-8p Sat 8a-5p, CT ***Pricing/Availability:** Pricing, specifications, availability, and terms of offer may change without notice. Taxes, fees, shipping, handling and any applicable restocking charges extra, vary and are not subject to discount. Offers may be combined with other select offers or discounts. U.S. Dell Small Business new purchases only. LIMIT 5 DISCOUNTED OR PROMOTIONAL ITEMS PER CUSTOMER. In case of customers leasing under these promotions, please note that items leased will be subject to applicable end-of-lease options or requirements. Dell cannot be responsible for pricing or other errors, and reserves the right to cancel orders arising from such errors. **Adobe Production Studio Standard:** Offer valid only with purchase of Dell Precision™ 380, 390, 490, 690, M65, M90. **Trademark/Copyright Notices:** Dell, the stylized E logo, E-Value, UltraSharp, CompleteCare and Dell Precision are trademarks of Dell Inc. Intel, Intel logo, Intel Inside, Intel Inside logo, Xeon, Xeon Inside, Intel Core, Core Inside are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. Microsoft and Windows are trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. Adobe, the Adobe Logo and Acrobat are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.

Related Categories:

Change management

Dell PowerEdge servers

Dell PowerVault storage

Firmware updates

Storage management

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

MAINTAINING DELL PLATFORMS WITH DELL TECHNICAL UPDATES

Dell periodically releases updates for its servers, storage, and other enterprise platforms. These updates are classified as urgent, recommended, or optional depending on the enhancements they provide. Dell recommends customers keep their system firmware and software up-to-date to take advantage of these enhancements, which are designed to improve system functionality and minimize potential problems. The following tables summarize recent firmware updates for Dell™ PowerEdge™ components and Dell PowerVault™ storage. To sign up for Dell Technical Updates, visit support.dell.com/support/notifications/technicalupdates.aspx.

Firmware updates for Dell PowerEdge Expandable RAID Controllers (PERCs), PowerVault direct attach storage, and hard disk drives

Product		PERC 3	PERC 4/DC, PERC 4/SC	PERC 4e/DC, PERC 4e/Di, PERC 4e/Si	PERC 4/Di	PowerVault 220S		Maxtor Atlas 10KV (Genesis)
Firmware version		199A	352B	522A	422A, 252A*	E.18	E.19	BP005
Firmware release date		September 11, 2006	September 11, 2006	September 11, 2006	September 11, 2006	October 12, 2005	April 18, 2006	December 8, 2006
Criticality		Urgent	Urgent	Urgent	Urgent	Urgent	Recommended	Urgent
Enhancement	Fixes problems with EVPD inquiry commands	■	■	■	■			
	Fixes write cache policy reporting	■	■	■	■			
	Fixes problem with cluster mode setting for multiple PERC 4/DC cards on a PowerEdge 6650		■					
	Reduces loss of storage access or communications					■		
	Minimizes SCSI resets, SCSI Enclosure Services time-outs, and MegaRAID (MRAID) errors					■		
	Improves SCSI parity						■	
	Provides Self-Monitoring, Analysis, and Reporting Technology (SMART) and minimizes time-outs							■

Firmware updates for Dell PowerVault tape drives

Product		PowerVault 110T LTO-2-L Certance	PowerVault 110T LTO-2	PowerVault 110T LTO-3	PowerVault 110T DLT VS160	PowerVault 100T DAT72	PowerVault 110T SDLT 320
Firmware version		1862, A12	53Y3, A04	5BG2, A01	3200, A13	A16E, A09	5D5D, A15
Firmware release date		June 19, 2006	April 17, 2006	April 18, 2006	November 20, 2006	November 16, 2005	April 10, 2006
Criticality		Urgent	Recommended	Recommended	Urgent	Recommended	Recommended
Enhancement	Improves load/unload operation	■	■	■	■	■	■
	Improves head cleaning	■				■	
	Improves error recovery	■	■	■	■	■	
	Reduces read/write errors	■	■	■	■	■	
	Reduces media failures	■	■	■	■	■	■
	Reduces eject/insert errors	■	■	■	■	■	■

Firmware updates for Dell PowerVault tape autoloaders and libraries

Product		PowerVault 122T LTO-2	PowerVault 124T LTO-2-L, LTO-3, and DLT VS160	PowerVault 132T and PowerVault 136T LTO-2	PowerVault 132T and PowerVault 136T SDLT 320
Firmware version		53Y3, A03	Loader V31	53Y3, A04	5D5D
Firmware release date		October 7, 2005	June 19, 2006	June 22, 2005	April 10, 2006
Criticality		Recommended	Recommended	Recommended	Recommended
Enhancement	Improves load/unload operation	■	■	■	■
	Improves head cleaning		■		
	Improves error recovery	■	■	■	
	Reduces read/write errors	■	■	■	
	Reduces media failures	■	■	■	■
	Reduces eject/insert errors	■	■	■	■
	Fixes move issues		■	■	
	Fixes communication issues		■	■	
	Fixes picker issues		■	■	

*Firmware update 422A applies to PowerEdge 1750 servers, and 252A applies to PowerEdge 2600 servers.

PURE DELL OPENMANAGE NEWSLETTER

Flexible Management for the Scalable Enterprise

February 2007

By Edward Reynolds, Senior Manager, Systems Management Product Marketing, Dell Inc.

View from the Top

Simplifying Client Management

Dell™ Client Manager™ software unifies several client administration tools into a single integrated solution that allows you to comprehensively manage your entire client environment—from hardware administration to software configuration and beyond—in a simple, scalable way. And it supports one crucial feature other client administration tools don't: secure remote-client reboots.

Client management applications have advanced over the years to include not only software configuration and hardware administration, but also advanced capabilities such as security policy compliance management, software license metering, and more. But there is one thing client management applications have not been able to do: reboot remote client systems in a secure way. And every time trained technicians leave their desks to reboot a client, productivity goes down the drain.

But times have changed. Dell Client Manager, the newly released client management application jointly developed by Dell and Altiris, supports the Alert Standard Format (ASF) 2.0 specification. Now, remote client computers that support ASF 2.0 can be power cycled in a secure way, helping eliminate the need to send valuable IT resources out on power-button-pushing pilgrimages.

Utilizing standards like ASF is just one of the many things Dell Client Manager does to help simplify client management and improve productivity. Dell Client Manager unifies the functionality of four client management tools—Dell OpenManage™ Client Instrumentation (the Dell client agent), the Dell Client Configuration Utility, Dell OpenManage IT Assistant, and Dell OpenManage Client Connector—into a single integrated solution that enables centralized, scalable management of complex client environments. And like other Dell hardware administration tools, Dell Client Manager Standard is available at no additional charge.

Dell Client Manager allows you to inventory and add computers to your client environment, configure and update BIOS images on multiple clients simultaneously, and monitor the health and status of key client computers, all from a single easy-to-use interface. Also, Dell Client Manager can be configured to filter and "listen" for alerts from critical systems—making it an excellent solution for monitoring mission-critical hardware such as ATM machines and point-of-sale clients.

For additional advanced client management capabilities, Dell Client Manager Standard can easily be upgraded to Dell Client Manager Plus, which extends the capabilities of Dell Client Manager to include deployment, basic software distribution, and PC migration. With Dell Client Manager Plus, you can capture and deploy system images, distribute and install software packages, migrate personal settings to new systems, and inventory system hardware and software information—all of which enable fast deployment of new systems, easy hardware upgrades, and simple user migrations.

But client management doesn't stop there. To help manage systems security and software license compliance, Dell collaborated with its longtime partner Altiris to offer Altiris® Client Management Suite™ software, which extends Dell Client Manager Plus by offering comprehensive policy-based functionality. Altiris Client Management Suite allows you to track the distribution and usage of software licenses, establish and enforce software configuration policies, and identify and remediate security issues with its patch management capabilities. IT administrators can control all of this from a centralized, easy-to-use console over a LAN, helping reduce the need for desk visits and speed up problem solving.

Dell Client Manager Standard, Dell Client Manager Plus, and Altiris Client Management Suite offer comprehensive, pay-as-you-grow functionality that allows you to automate and simplify the management of your client environment. Whatever the size of your organization, Dell Client Manager can help you reduce costs by streamlining IT administration and enhancing the productivity of your IT staff.

Oh, and it supports secure remote-client reboots.

INSIDE THIS ISSUE:

Law School Invokes Client Management Privileges

Northwestern University School of Law simplifies client management with Dell™ Client Manager™ Plus and Altiris® Client Management Suite™ software

Located in the heart of downtown Chicago, Northwestern University School of Law is a top-tier U.S. law school. With nearly 1,000 students and 350 faculty and staff, it educates some of the best and brightest in the nation. At Northwestern, the burgeoning legal minds are embracing the use of technology in what has previously been a very traditional profession. "We want technology to be a tool for our students, but not their primary focus," says Christel Bridges, IT director at Northwestern University School of Law. "We want its use to be both ubiquitous and transparent."

Deploying and supporting multiple clients—including Dell™ Latitude™ notebooks and Dell OptiPlex™ desktops—across the urban campus has been a challenge for the eight-person IT team. Daily client maintenance practically required the IT department at the law school to issue subpoenas to track down its IT personnel. "Any time we needed to inventory clients—to check memory, hard disk space, or application profiles—we had to physically touch the equipment," says Bridges. "Also, our clients were becoming increasingly mobile, so scheduling time to troubleshoot problems drained IT personnel time and inconvenienced users."

Northwestern University School of Law attempted to find a client management tool that could perform maintenance and inventory remotely, but the products the IT team tried were too difficult to use. "We couldn't afford to have a person dedicated to learning the management tool. We needed to simplify the process," says Bridges. "At a technology conference, I heard about a client management tool from Altiris that had normally cynical IT people practically gushing about its features and ease of use."

Because of the Dell partnership with Altiris, Bridges looked to Dell to help implement a client management solution. "A consultant from Dell Deployment Services helped us deploy and integrate our management tools," says Ping Ng, an IT administrator at the law school. "We were also using Dell OpenManage™ software to help maintain our Dell PowerEdge™ servers, so the consultant was able to help us integrate tools from Dell and Altiris. And since they have a strong partnership, all the management consoles are consistent, which makes training easy."

Using Altiris Client Management Suite, Northwestern University School of Law is enabling its small IT team to provide unobtrusive client management and support while simplifying common IT tasks. "Being able to remotely update the BIOS of our Dell OptiPlex desktops and Dell Latitude notebooks has really saved us from running around and tracking these systems down," says Ng. "And the full imaging that Altiris software can do has reduced our new hardware deployment time from two hours to 45 minutes."

Bridges echoes Ng's sentiments from a management perspective: "The inventory capabilities of our management tools enable us to track everything on our clients, from hardware components to OS versions. I can provide detailed reports of our entire inventory to university administrators, including warranty information for hundreds of systems."

"With Altiris Client Management Suite, we have reduced the time spent troubleshooting and imaging clients by half—enabling us to realign our resources to other IT projects that can improve the educational experience of our students and faculty."

—Christel Bridges, IT Director
Northwestern University
School of Law
June 2006

The Northwestern University School of Law IT team now spends less time mediating client problems. "Using Altiris Client Management Suite, we have reduced the time spent troubleshooting and imaging clients by half," says Bridges. "That has enabled us to realign our resources to other IT projects that can improve the educational experience."

Challenge

Inefficient client management tools forced small IT team to physically track down notebooks and visit desktops to inventory hardware and software; client image rebuilding drained limited IT resources

Solution

Hardware:
- Dell Latitude notebooks
- Dell OptiPlex desktops
- Dell PowerEdge servers

Software:
- Altiris Client Management Suite
- Dell OpenManage Server Administrator

Services:
- Dell Deployment Services

Benefits

- Management tools provide a centralized, easy-to-use console that helps the IT team simplify server and client management
- Remote management capabilities enable the IT team to perform BIOS updates and troubleshooting in half the time previously spent on such tasks
- Client management tools provide the IT team with a comprehensive inventory of hardware and software assets
- Full client imaging capabilities help reduce new hardware deployment from two hours to 45 minutes

TECH CORNER

By James Lathan, Technology Specialist, Dell Inc.

Taking Control of Your Client Environment

Administering far-flung client systems without the right tools can be time-consuming and problematic. Now, Dell™ Client Manager™ and Altiris® Client Management Suite™ software offer an integrated, easy-to-use management framework for proactively maintaining your client environment—helping contain costs, increase user productivity, and anticipate problems before they occur.

Problem: *How can our IT team ensure a consistent client image across the organization?*

Consistency should start with the BIOS revisions and settings. For example, Dell Client Manager Standard manages the BIOS settings of Dell OptiPlex™ desktops, Dell Latitude™ notebooks, and Dell Precision™ workstations on multiple client systems to help ensure system consistency with aspects of the boot sequence. Many organizations that support a varied environment resist fully enforcing consistency. However, using a BIOS image that is consistent across an organization, or even just across a department, can reduce troubleshooting steps for the help desk and help lower problem remediation time.

Of course, BIOS consistency is just the start. Dell Client Manager Plus extends the management infrastructure to client software, allowing administrators to distribute consistent software packages and configurations across an organization to help simplify new client provisioning and reduce troubleshooting complexity. Not only does Dell Client Manager Plus update and distribute Dell BIOS settings, it also enables administrators to set and quickly deploy images for different departments.

Problem: *How can our IT team stay ahead of costly and time-consuming hardware failures?*

Proactively monitoring the state of client hardware with Dell Client Manager Standard helps avoid downtime and data loss, which may compromise end-user productivity. For example, by monitoring the health of a hard drive, Dell Client Manager Standard can notify administrators when a technician should be dispatched quickly to remediate an issue. Previous Dell client management tools monitored only a handful of alerts, but the latest release of Dell Client Manager Standard enables administrators to monitor 17 different client status alerts.

If administrators receive a pre-failure alert from a client system, they can automate backup by executing a command on the client system to either shut down or perform a backup to help prevent data loss and reduce downtime. Also, Dell Client Manager Plus enables administrators to capture user settings and files, which streamlines the process of upgrading existing systems or migrating users to new systems.

Problem: *How can our IT team keep track of asset inventory from a central location—not just devices, but also software and integrated hardware components?*

Dell Client Manager Standard is designed to perform a detailed hardware inventory and store the systems inventory in the Dell Client Manager database. For example, by using the inventory capabilities of Dell Client Manager Standard, administrators could determine the battery manufacturer of client systems—enabling many organizations to eliminate thousands, if not all, of their clients from the scope of a recent battery recall. Knowing exactly which components are in the IT environment helps IT staffers avoid spending time on problems that do not apply to their systems.

Using Dell Client Manager Plus, IT departments can extend the inventory capabilities to the software level of their client systems. Storing detailed OS and application information in the Dell Client Manager database and combining it with an inventory of hardware details can provide IT departments with a comprehensive view of their client environments. That knowledge, combined with the ability to configure certain aspects of the clients, provides a level of control that eludes many IT organizations today.

Problem: *How can our IT team begin to plan a large-scale client OS upgrade when we do not know if our hardware can accommodate the performance requirements of the new OS?*

By starting with the inventory capabilities of Dell Client Manager Standard and Dell Client Manager Plus, administrators can obtain a comprehensive snapshot of their client environments. They can identify the memory level of clients, graphics card models, and current OS version levels. For example, by leveraging the Readiness Reports within Dell Client Manager Standard and Dell Client Manager Plus, administrators can determine which systems are capable and which need to be replaced to run the Microsoft® Windows Vista™ OS. Further analysis of this data can guide IT administrators in determining which migration options may be cost-effective for their organization.

You may not gain complete control over a rapidly expanding and evolving client environment overnight, but Dell Client Manager Standard helps IT departments maintain a complete inventory of their systems, proactively monitor client health, and provide consistency across clients. After these initial steps, enhanced client control is available by using the software management capabilities of Dell Client Manager Plus and the security policy compliance capabilities of Altiris Client Management Suite. Instead of reacting to client management issues as they arise, you can take advantage of the appropriate tools to streamline IT operations and help ensure the cost-effective utilization of IT resources. ■

The inventory capabilities of Dell Client Manager Standard enabled many organizations to eliminate thousands of clients from the scope of a recent battery recall.

PARTNER INSIGHTS

By Steve Morton, Vice President of Product Management and Marketing, Altiris, Inc.

Altiris and Dell: Architecting the Future of IT Management

Dell and Altiris share the vision of simplified management by offering standards-based products that are flexible, interoperable, and scalable. The Dell Scalable Enterprise Reference Architecture codifies this vision into a set of guidelines for designing, building, and assessing the merits of IT components. Altiris and Dell have partnered to put this vision into action with a comprehensive range of management applications, including the Dell™ Client Manager™ suite.

As enterprise data centers expand to meet emerging business and operational challenges, so does the complexity of managing them. To help IT departments deliver increasingly sophisticated functionality in a manageable way, Dell has developed the Dell Scalable Enterprise Reference Architecture, a framework for designing and implementing efficient, manageable IT infrastructures that have the inherent flexibility and scalability to help meet the changing needs of today's businesses.

At its heart, the Dell Scalable Enterprise Reference Architecture recommends that data centers be built using scalable, standards-based components that can be flexibly combined to deliver enhanced business functionality. Altiris® products are built on the principles articulated in the Dell Scalable Enterprise Reference Architecture.

Consider, for example, the Dell Client Manager suite of client management applications, jointly developed by Altiris and Dell. The Dell Client Manager family consists of Dell Client Manager Standard, which provides centralized hardware administration; Dell Client Manager Plus, which adds software deployment and configuration capabilities; and Altiris Client Management Suite™ software, which offers comprehensive client administration that includes advanced features such as security compliance and policy management.

Because Altiris products, including the Dell Client Manager suite, are built on a common architecture and share a common data model, they can either be deployed as stand-alone solutions or easily combined to deliver more extensive functionality than each could provide individually—giving customers the flexibility to choose only those products they need and the scalability to add functionality as business requirements evolve.

Furthermore, because the Dell Client Manager family adheres to industry-recognized standards such as the Distributed Management Task Force's Common Information Model (CIM) and the Windows Management Instrumentation (WMI) standard, they can easily be used in concert with other systems management tools, such as Microsoft® Systems Management Server. And because Altiris products leverage industry-standard Internet protocols such as HTTP, HTTP over Secure Sockets Layer (HTTPS), and XML, comprehensive functionality can be accessed securely and easily through a standard Web browser.

At Altiris, we understand that the data center of the future will need to leverage standards-based, scalable components to deliver increasingly sophisticated business functionality in an efficient, manageable way. So does Dell—which is why we have committed to jointly develop all Dell-branded console applications.

For more information about the alignment between the Dell scalable enterprise vision and Altiris architecture, visit www.dell.com/content/topics/global.aspx/vectors/en/2006_altiris_scalableentref.

February 2007

Servers

Storage

Systems Management

Services

Software

ORACLE®

How Dell Does IT

Dell™ understands that mission critical databases require powerful performance, highly reliable and available hardware platforms, and scalability that allows you to grow with your database needs. Dell understands this so well we run our own supply chain management system, which supports hundreds of suppliers and manages nearly a billion parts per year, on powerful and scalable Dell PowerEdge™ servers running Oracle® Database 10*g*. Dell offers pre-engineered, tested, and validated solutions for Oracle Database 10*g* on Linux® and Windows®. The entire solution stack is tested and supported by Dell – the servers, storage, switches, and software, including the Oracle Database software and the operating system – and Dell offers services to help you accelerate deployment.

Go to dell.com/oraclemag to get the full story on how Dell IT uses Oracle Database 10*g* in our supply chain management system.

Exploring the Remote Access Configuration Utility

in Ninth-Generation Dell PowerEdge Servers

The Remote Access Configuration Utility supports local and remote server management in ninth-generation Dell™ PowerEdge™ servers using the Intelligent Platform Management Interface (IPMI); it combines the features of the eighth-generation Remote Access Configuration Utility and IPMI Server Management Configuration Utility and introduces several important features. This article describes the feature set and configuration settings provided by this utility.

BY KALYANI KHOBRAGADE

Related Categories:

Baseboard management controller (BMC)

Dell ninth-generation servers

Remote access controllers (RACs)

Systems management

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

As systems management becomes more complex, it is increasingly important for administrators to be able to quickly monitor and manage their servers. Dell PowerEdge servers include a baseboard management controller (BMC) designed to help administrators proactively monitor and manage server hardware, log server fault events, and alert administrators when server faults occur. The BMC also provides remote capabilities such as changing the power state of the server and Intelligent Platform Management Interface (IPMI) Over LAN. The Dell Remote Access Controller (DRAC), an optional feature for PowerEdge servers designed to allow administrators to remotely configure and manage servers, works in conjunction with the BMC to enable advanced remote management features such as virtual media, dedicated network controllers, and remote consoles.

The Remote Access Configuration Utility, also known as the BMC option ROM, resides in the BIOS and provides a simple user interface engine that allows BMC and DRAC configuration through an integrated set of rich systems management features. This utility combines the features of the eighth-generation Remote Access Configuration Utility and IPMI Server Management Configuration Utility and has been extensively augmented with enhanced features. It launches during the system power-on self-test (POST) in a non-OS-based environment, thus removing dependency on other utilities like the Dell OpenManage™ suite and enabling quick BMC and DRAC configuration. Because the utility is embedded in the system BIOS, no separate installation is required. This utility can be particularly useful to organizations that require configuration capabilities but not a complete set of management tools.

Administrators can launch the utility by pressing Ctrl + E during the system BIOS option ROM scan. The utility supports password protection and remote configuration using console redirection.

Utility launch

The system BIOS loads the utility as a standard option ROM during the POST, displaying and allowing administrators to

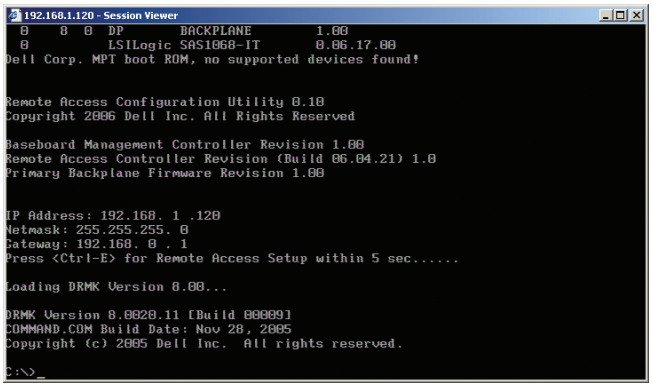


Figure 1. Remote Access Configuration Utility during the POST

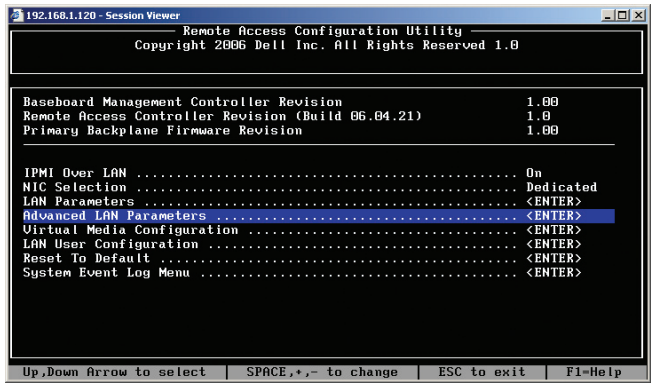


Figure 2. Remote Access Configuration Utility setup module

configure BMC and DRAC settings. The utility loads the initialization module, which displays the BMC banner and prompts the administrator to enter the setup module by pressing Ctrl + E within 5 seconds. Administrators can skip the five-second wait period and the setup module by pressing Esc or the space bar. If the administrator does not press Ctrl + E within five seconds, the POST automatically skips the setup module.

Initialization module

By default, the initialization module performs BMC initialization tasks and displays BMC information in the BMC banner. After establishing communication with the firmware, it also displays the BMC communication failure check; the BMC, DRAC, backplane, and flex bay versions; the cable errors check; and emergency management port settings. Figure 1 shows the information displayed by the initialization module during the POST.

BMC communication failure check

During initialization, the utility attempts to determine whether BMC communication is reliable by issuing IPMI communication. If this communication fails, the utility displays the message "BMC Communication Failure" and does not provide the Ctrl + E launch option.

BMC, DRAC, backplane, and flex bay versions

If the IPMI communication with the BMC is successful, the utility displays the BMC firmware version by default. If it detects a DRAC, backplane, or flex bay, the BMC tries to communicate with them; on successful communication, the utility displays the corresponding firmware revisions. If the backplane or flex bay are in boot block mode or the versions do not match and must be upgraded to the latest version, the utility displays the warning message "Backplane firmware is out-of-date. Please update to the latest firmware."

Cable errors check

Serial Attached SCSI (SAS) cables should be connected using a valid SAS cable configuration. Otherwise, the initialization module displays a separate message for each cable error—for example, "SAS A Cable Missing/Misconfigured," "SAS B Cable Missing/Misconfigured," "Flex Bay Cable Missing/Misconfigured," or "Flex Bay Power Cable Missing/Misconfigured." It also displays the message "Power down the system and connect the cable correctly. Details for storage cabling can be found on the system information label and in the Hardware Owner's Manual."

Emergency management port settings

This section of the utility displays the current LAN parameter settings. By default, the initialization module displays the IP address, subnet mask, and gateway settings as zeros unless the administrator configures these parameters after pressing Ctrl + E. The module then displays the updated settings during the POST on subsequent reboots.

Setup module

Administrators can launch the setup module during the prompt "Press <Ctrl-E> for Remote Access Setup within 5 sec." During this time, a period is appended to the end of the prompt text each second. Pressing Ctrl + E invokes the setup module (see Figure 2), which displays BMC and DRAC settings and allows administrators to configure these settings through the user interface engine (see Figure 3).

The following sections describe the IPMI Over LAN, NIC Selection, LAN Parameters, LAN User Configuration, and Reset To Default settings. The "Feature enhancements in ninth-generation Dell PowerEdge servers" section in this article discusses the Advanced LAN Parameters, Virtual Media Configuration, and System Event Log Menu settings.

IPMI Over LAN

Using IPMI Over LAN requires the managed node to be connected to the management station using a LAN cable on the same subnet, thus allowing administrators to perform remote server management. Administrators should turn the IPMI Over LAN option on before setting any other LAN options. Turning this option off disables all LAN-specific options.

NIC Selection

The NIC Selection option includes three settings: Shared, Failover, and Dedicated. If a DRAC is not present or does not respond to BMC commands, the Dedicated setting is hidden; if a DRAC is present, this field uses the Dedicated setting by default. When a DRAC is present, all network communication is managed by the DRAC instead of the BMC.

Ninth-generation Dell servers have two embedded network interface cards (NICs). Administrators can use these NICs for server management in two configurations based on the Shared and Failover settings. The Shared setting configures the BMC or DRAC to transmit and receive network packets using NIC1 and only receive packets using NIC2. If NIC1 fails, then the remote management station cannot communicate with the server. Administrators may want to choose this setting if they are using only one NIC for server management.

The Failover setting provides communication over both NICs. Administrators can choose this setting when using both NICs for server management and when they want server management connectivity even if one NIC fails. In this mode, the BMC or DRAC receives network packets using NIC1 and NIC2, but only transmits packets using NIC1. If NIC1 fails, then NIC2 begins transmitting and receiving; when NIC1 resumes, it will only receive. If NIC2 then fails, NIC1 begins transmitting and receiving again, and when NIC2 resumes, it will only receive.

LAN Parameters

The LAN Parameters settings include RMCP+ Encryption Key, IP Address Source, Ethernet IP Address, MAC Address, Subnet Mask, Default Gateway, VLAN Enable, VLAN ID, VLAN, LAN Alert Enabled, Alert Policy Entry 1, Alert Destination 1, and Host Name String.

RMCP+ Encryption Key. The IPMI 2.0 enhanced authentication option enables administrators to connect to the server using a Remote Management Control Protocol+ (RMCP+) key. If the key is specified in this field, administrators must provide the key to communicate remotely with the BMC and DRAC.

IP Address Source. Before a computer connects to a network, it must obtain an address that uniquely identifies it—either a Dynamic Host Configuration Protocol (DHCP) or a configurable (static) IP address. This field is read during the POST and updated whenever administrators choose the Reset To Default option. The default settings are initialized to "Unknown" and allow administrators to select the desired source.

Ethernet IP Address, MAC Address, Subnet Mask, and Default Gateway. Administrators must configure the IP address, subnet mask, and gateway according to their network requirements and settings. Certain values are invalid under IP rules—for example, the utility limits addresses to values under 255.255.255.255. IP addresses must also comply with the following rules:

Option	Settings	
IPMI Over LAN	On/Off	
NIC Selection	Shared/Failover/Dedicated	
LAN Parameters	RMCP+ Encryption Key IP Address Source (Static/DHCP) Ethernet IP Address MAC Address Subnet Mask Default Gateway VLAN Enable VLAN ID (0–4095) VLAN (Priority 0–7) LAN Alert Enabled Alert Policy Entry 1 Alert Destination 1 Host Name String	
Advanced LAN Parameters	Dedicated NIC Configuration Options	NIC (Enabled/Disabled) Auto-Negotiate (Enabled/Disabled) LAN Speed Setting (10 Mbps/100 Mbps) LAN Duplex Setting (Full Duplex/Half Duplex)
	DNS Configuration Options	DNS Servers from DHCP (On/Off) DNS Server 1 DNS Server 2 Register RAC Name (On/Off) Register RAC Name (up to 32 characters) Domain Name from DHCP (On/Off) Domain Name (up to 64 characters)
Virtual Media Configuration	Virtual Media (Attach/Detach) Virtual Flash (Enable/Disable)	
LAN User Configuration	Account Access Account Privilege Account User Name Enter Password Confirm Password	
System Event Log Menu	Total System Event Log Entries View System Event Log Clear System Event Log	

Figure 3. Remote Access Configuration Utility setup module options

- The first octet must be between 1 and 223, but cannot be 127 (for example, 143.*xxx.xxx.xxx*, where *xxx* is a number between 0 and 255).
- The last octet must not be 0 or 255 (that is, *xxx.xxx.xxx*.0 or *xxx.xxx.xxx*.255).

The MAC Address field is a read-only field showing the Media Access Control (MAC) address for the BMC NIC.

Remember when attacks occurred on the outside?
Today, more and more attacks are coming from within your network.

Do you have the right security solution?

The Viking V 2-Gigabit LAN Firewall

The Viking V 2-Gigabit LAN Firewall offers premium performance at a fraction of the price. It's the only firewall that performs at full Gigabit speeds in its price range – even when fully configured.

And independent tests confirm it!*

Traditional firewall use at the network's perimeter protects from outside attacks, but today more and more attacks are occurring from the inside. Is your LAN protected?

Learn more at: www.vikinginterworks.com/security

VIKING INTERWORKS™

Tel: +1 408 964 3947
Tel: 800 529 7710 (US and Canada only)
Fax: +1 408 964 3437
Email: security@vikinginterworks.com

*Up to Spec is The Tolly Group's flagship custom testing program for certification of Information Technology products, services and components. The Tolly Group offers uncompromised, independent evaluations.

VLAN Enable, VLAN ID, and VLAN. The virtual LAN (VLAN) fields allow the creation of independent logical networks within a physical network, helping simplify the creation of sub-networks. The settings should conform to the IEEE 802.1Q specification. The ability to move clients to different broadcast domains by setting membership profiles for each port on centrally managed switches is one of the main advantages of 802.1Q VLANs. Although management stations and servers can be located anywhere on a network, they are grouped together by VLAN technology, and broadcasts are sent to devices within the VLAN.

LAN Alert Enabled, Alert Policy Entry 1, Alert Destination 1, and Host Name String. Because of the possibly critical nature of alerts, administrators can configure servers to send Simple Network Management Protocol (SNMP) traps to a remote destination. Administrators can configure alert settings to indicate sensor states to the client system (the management station) specified in the Alert Destination 1 field. Alert severities include critical, warning, and informational. Traps are sent to this destination only when the LAN Alert Enabled setting is turned on.

Platform Event Traps (PETs) originating from the firmware may contain up to 48 bytes of an original equipment manufacturer (OEM) ASCII string appended to the OEM section of the PET. This string typically contains a unique description of the host server contained in the Host Name String setting.

LAN User Configuration

The LAN User Configuration settings include Account Access, Account Privilege, Account User Name, Enter Password, and Confirm Password.

Account Access and Account Privilege. These fields allow administrators to permit specific users to access the servers with a defined privilege level. Users can take actions remotely based on their privilege settings.

Account User Name, Enter Password, and Confirm Password. Administrators can change the username and password for the default administrator account, UserID2, as defined by the IPMI specification. The utility assumes this account will always be enabled and have administrative privileges. A successful password change requires the administrator to reenter the password for confirmation. If the administrator does not reenter the password correctly, the password will not be changed.

Reset To Default

Dell firmware supports resetting certain nonvolatile settings to the factory defaults. No password or authentication is required to issue the Reset To Default command, because all local communication is assumed to have administrative privileges. To protect against accidental resetting, the utility provides a cancel option as the default state.

Feature enhancements in ninth-generation Dell PowerEdge servers

In addition to the features described in the "Setup module" section in this article, the Remote Access Configuration Utility includes enhanced capabilities such as password authentication, a DRAC feature set (the Advanced LAN Parameters and Virtual Media Configuration settings), and the System Event Log Menu settings.

Password authentication

The entire setup module can be password protected for enhanced security. The setup password cannot exceed 32 characters. If administrators have enabled this feature in the BIOS, the utility queries for the password. Following successful password verification, the administrator has access to all fields; if the password is not entered successfully in three attempts, the utility limits all fields to read-only status. *Note:* Once in this read-only mode, the utility cannot query for the password again. If administrators want to enter the password to gain full access, they must reboot the server.

If the administrator presses the Esc key during password entry, all fields are disabled. Administrators cannot change the password within the utility. Once disabled, the field values are visible but not changeable.

DRAC feature set

For ninth-generation Dell PowerEdge servers, the DRAC no longer resides on the PCI bus. Because of this change, ninth-generation Dell server BIOSs no longer include the DRAC option ROM Ctrl + D option to read and configure DRAC settings. Instead, administrators can carry out DRAC configuration using the BMC option ROM.

During initialization, the BMC tries to establish a connection to the DRAC. If the DRAC does not respond to BMC commands within three tries, the BMC considers the DRAC missing and hides all DRAC-specific components in the utility. When a DRAC is detected, the utility exposes the Advanced LAN Parameters and Virtual Media Configuration settings (see Figure 4).

Advanced LAN Parameters. Administrators should keep in mind that the NIC integrated in the DRAC cannot be used for OS network traffic; it can be used only to communicate with the DRAC to perform management functions.

Based on the DRAC presence signal, the BMC reads the DRAC revision, NIC settings, and Domain Name System (DNS) settings. Any changes are saved to the DRAC when the administrator saves and exits. If the commands fail, the utility continues the POST but displays the message "BMC Communication Error." Remote users are allowed to communicate with the DRAC based on these settings.

Virtual Media Configuration. The Virtual Media Configuration option, which includes the Virtual Media and Virtual Flash features, helps administrators use data center resources efficiently and maximize the power and manageability of server environments.

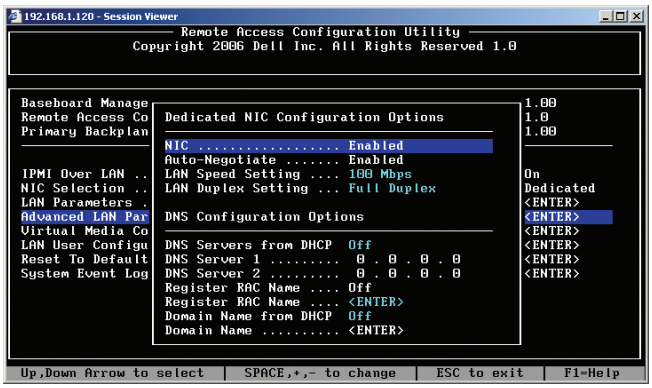


Figure 4. DRAC Advanced LAN Parameters settings in the Remote Access Configuration Utility

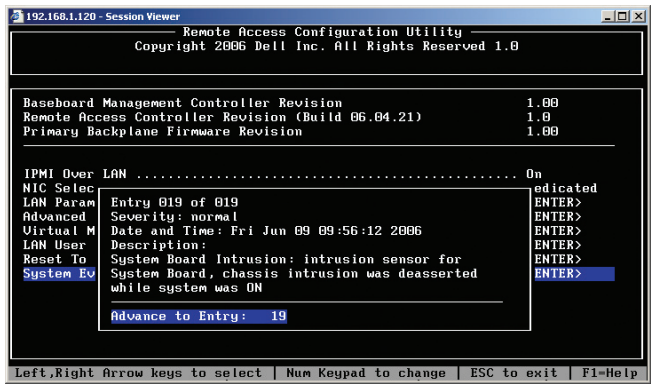


Figure 5. Remote Access Configuration Utility System Event Log Menu showing the most recent log entry

Administrators can control the OS and BIOS of servers remotely using virtual media access.

The Virtual Media feature creates a connection between storage media devices on the client system (the management station) and the target platform (Dell PowerEdge servers). This feature is particularly useful when remotely installing BIOS and firmware or OS components. The DRAC virtual media capabilities are designed to enable organizations to access remote media such as CDs, DVDs, ISO images and floppy disks, USB keys, and floppy images through the host as though they were physically present on the system. The server identifies these drives during OS boot and after the OS is completely booted.

The options available under the Virtual Media setting are Attach and Detach. Because the DRAC in ninth-generation Dell servers (DRAC 5) supports virtual media implementation based on USB, it supports seamless attach and detach operations for virtual media devices. The Attach option attaches virtual media devices to the USB bus, after which the devices are available to administrators. The Detach option removes all virtual media devices from the USB bus.

The DRAC 5 also supports virtual flash (USB key) in its virtual media implementation. The Virtual Flash feature is particularly useful because it does not require external client connections or devices to be functional in the host server. Currently the virtual flash size is 16 MB; it appears as an unformatted removable USB drive in the host environment. The USB flash is disabled by default. The Virtual Flash feature allows virtual flash media to be attached to the server as if it were physically located on the server. The Enable option allows the virtual flash to be available for use if virtual media is attached. The Disable option causes the virtual flash to be removed and makes it unavailable for use. *Note:* Enabling and disabling this feature causes all virtual media devices to be detached from and attached to the USB bus, meaning that the detach and attach operations interfere with data read/write operations from virtual media devices on the server.

System Event Log Menu

The system event log captures critical server events. Log entries are time-stamped and contain the severity and description of an event. The Remote Access Configuration Utility allows administrators to read and clear the log. It also implements the common text format to display the log as used by other Dell programs and utilities.

When the administrator selects the System Event Log Menu option and presses the Enter key, the utility displays a progress bar while loading the Sensor Data Records (SDRs). If the system does not load the SDRs properly, it cannot display the system event log. If it loads the SDRs successfully, the default view shows three fields: Total System Event Log Entries, View System Event Log, and Clear System Event Log.

The System Event Log Menu displays the number of events present in the log. Administrators can view the log using the View System Event Log field. By default, the utility displays the most recent log entry (see Figure 5). Administrators can use the arrow keys to navigate through the log entries or view a specific entry by specifying the entry number. The utility only allows a log index range between 1 and the total number of log events; thus, log indexes less than 1 and greater than 512 are ignored.

Streamlined systems management

The Remote Access Configuration Utility enables administrators to view and configure local and remote systems management settings in Dell server environments. To supplement the utility's help text available using the F1 key, administrators can use this article as a guide for optimizing BMC and DRAC configuration settings.

Kalyani Khobragade is a development engineer in the Dell Remote Management Firmware team. She was previously on the Dell BMC team. Her interests include digital signal processing and wireless communications. She has a B.S. in Electronics Engineering and an M.S. in Electronics Engineering from Visvesvaraya Regional College of Engineering, and an M.S. in Electrical and Computer Engineering from Purdue University.

Dell IT Designs an Enterprise Monitoring Infrastructure

Using Microsoft Operations Manager 2005

Microsoft® Operations Manager (MOM) 2005 provides scalability, failover, integration with Dell OpenManage™ Server Administrator, and other features enabling it to serve as the foundation of a global Microsoft Windows Server® OS–based enterprise monitoring infrastructure. Using MOM 2005 along with Dell™ PowerEdge™ servers, Dell PowerVault™ storage, Dell OpenManage Server Administrator, and BMC software, the Dell IT group developed a robust infrastructure to support more than 24,000 servers worldwide.

BY DAVID RUBIO, LON ALONZO, AND PRAGNESH RATHOD

Related Categories:

Case study

Dell PowerEdge servers

Enterprise management

Microsoft Operations Manager (MOM)

Scalable enterprise

Storage management

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

When the Dell IT group, after years of supporting increasing revenue and unit shipments, assessed its environment to determine how to streamline server monitoring, it recognized the potential benefits of introducing additional centralization and standardization. Its goal was ensuring server stability across the enterprise while maximizing the return on investment in the monitoring infrastructure. After considering a number of possibilities, Dell IT ultimately chose a system based on Microsoft Operations Manager (MOM) 2005.

Global monitoring challenges

Initial Dell IT assessments of its monitoring environment revealed several problems with the existing infrastructure:

- The various environments, which involved numerous systems and products, could not scale to support the entire enterprise, and a transition to a centralized model would require tripling the largest monitoring system's capacity.
- The agent health monitoring was not robust enough to alert support personnel when server monitoring on specific agents failed.
- The environment scale, involving tens of thousands of servers spread across the globe, increased the difficulty of applying consistent systems and practices.

After assessing these problems, Dell IT determined that to implement an efficient global infrastructure, it would need to create consistency in monitoring standards, design a solution capable of accommodating complex support groups and processes with seemingly contradictory requirements, and integrate the numerous server monitoring systems then in use. Dell IT subsequently conducted a

formal product evaluation and identified several requirements for a proposed monitoring application, including the following:

- **Scalability:** The application would need to scale well to monitor tens of thousands of servers. An application requiring a large, costly, complex infrastructure or incapable of delivering standardized rules to large groups of servers could become too unwieldy for support teams to manage, increasing the risk of system instability.
- **Extensive feature set:** The application would need to support hardware failure detection, monitoring with Microsoft Windows® OS event logs and Performance Monitor, monitoring of common Microsoft Windows Server OS–based applications, custom monitoring with common provider interfaces such as Windows Management Instrumentation (WMI), and robust agent health features to alert support teams when monitoring on critical servers fails. Discovery capabilities were also essential, to help ensure comprehensive server coverage and reduce the need for human involvement when beginning monitoring.
- **High availability:** The application would need to support high-availability features such as clustering and agent failover, so that support teams could still monitor revenue-affecting outages following a failure.
- **Flexibility and ease of use:** The application would need to be flexible enough to accommodate individual exceptions to common thresholds or parameters, and possess customizable interfaces and consoles for distribution to regional support teams. Ease of use was necessary to help minimize operational overhead.

Infrastructure foundation: MOM 2005

Dell IT selected MOM 2005 as the foundation of its monitoring infrastructure because it met the most requirements of any candidate product. In terms of scalability, for example, a five-server MOM management group could support up to 4,000 monitored servers, which would allow a relatively small number of management groups to manage the global Windows Server environment. Dell IT also added extra management groups to help support disaster recovery and accommodate future growth.

MOM 2005 possesses a number of other important features. With the addition of the Dell management pack (MP) and Dell OpenManage Server Administrator agents on Dell PowerEdge servers, it can provide extensive information on hardware events such as drive and fan failures, memory errors, and temperature problems, and administrators can calibrate and suppress alerts to provide support groups with an appropriate level of instrumentation. MOM provides straightforward monitoring with Windows event logs and Performance Monitor, and enables administrators to use scripts and compiled code for custom monitoring using providers such as WMI as well as Microsoft Visual Basic®, Microsoft Visual Basic .NET, C#, and other languages. MOM also supports high-availability features such as agent failover and Microsoft SQL Server™ clustering.

Feature	Previous infrastructure	MOM 2005–based infrastructure
Maximum number of managed agents	5,000	More than 24,000
Servers required to support maximum capacity	76	24
Attributes-based rule deployment	No	Yes
Global reporting	No	Yes
Global standards	No	Yes
Automated deployment based on asset management database	No	Yes
Monitoring and performance data encryption	No	Yes

Figure 1. Comparison of previous Dell IT monitoring infrastructure with Microsoft Operations Manager 2005–based infrastructure

In addition, MOM can provide the flexibility and ease of use Dell IT requires. MOM MPs for a wide variety of applications, including SQL Server 2005, Microsoft Exchange Server 2003, Microsoft SharePoint® Portal Server 2003, Microsoft Windows SharePoint Services 2003, and Citrix MetaFrame XP, allow Dell IT to extend its investment in MOM to other applications in the future. And the MOM Administrator and Operator consoles enable regional Dell IT teams to create custom views showing only the servers, alerts, and severities relevant to their role. Figure 1 summarizes some of the advantages the MOM 2005–based infrastructure offers over the previous Dell IT monitoring infrastructure.

Management groups

The Dell IT monitoring infrastructure is based on MOM 2005 management groups, which consist of a MOM database and one or more MOM management servers (see Figure 2). Each management group is a distinct entity that is unaware of other management groups. Although multitiered options exist that can unify multiple management groups, the Dell IT requirements would test the sizing guidelines that make these options feasible.

Dell IT created three types of MOM management groups:

- **Production:** These groups monitor production systems and are configured with high-availability features and the stringent parameters and thresholds required by Dell production

environments. Dell IT chose a five-server design for these management groups to help provide high availability along with sufficient capacity to support a management server failure.

- **Disaster recovery:** These groups are designed to provide minimally sufficient monitoring if an entire production data center or management group fails or becomes unavailable; they do not contain high-availability features, but do contain the same rules and thresholds as production server management groups. Including limited capabilities in these management groups helps reduce overall infrastructure costs.
- **Non-production:** These groups manage non-production environments such as development and testing servers, which Dell IT does not manage as rigorously as production environments. Including reduced high-availability features and lower parameters and thresholds enables support teams to appropriately prioritize alerts from these environments, helping reduce overall infrastructure costs.

Each production management group includes three management servers and a cluster of two database servers. The three production management servers are Dell PowerEdge 2850 servers running Windows Server 2003 Enterprise Edition. These servers comply with MOM sizing guidelines to support 2,000 agents each; by allotting three management servers to support a maximum of 4,000 agents, the management group can use the MOM management server failover feature to accommodate the failure of one management server without affecting performance.

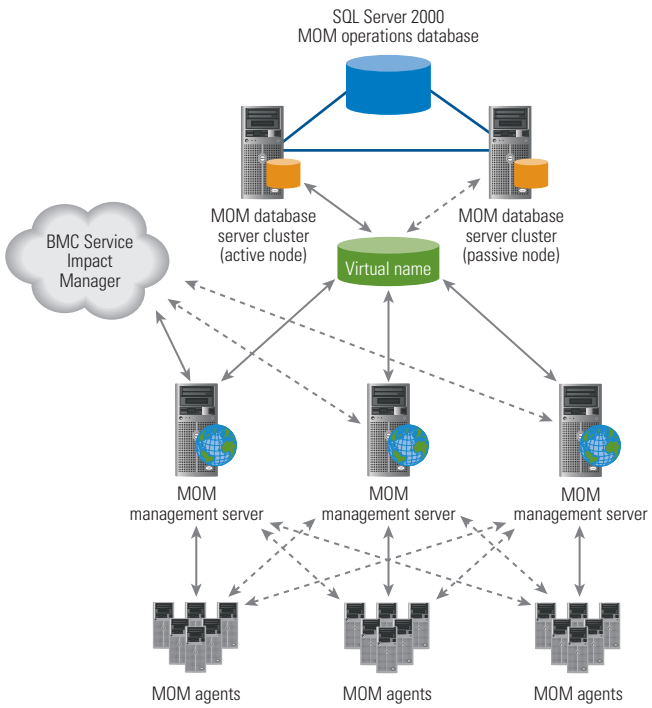


Figure 2. Dell IT monitoring infrastructure based on Microsoft Operations Manager 2005 management groups

The MOM database is hosted by a cluster of two Dell PowerEdge 6850 servers running Windows Server 2003 Enterprise Edition clustering services. These two servers function in an active/passive configuration to help ensure that the MOM database is available even in the unlikely event of a complete server failure, and they share a storage area network (SAN) based on Dell PowerVault storage to help provide optimum database read/write performance.

Figure 3 shows the global Dell IT monitoring infrastructure deployment, which consists of two production, two disaster recovery, and two non-production management groups in the Americas region; one production and one disaster recovery management group in the Asia Pacific and Japan region; and one production and one disaster recovery management group in the Europe and Middle East region.

Management packs

A substantial part of implementing the Dell IT monitoring infrastructure consisted of globally standardizing OS and server hardware monitoring. Two primary MOM 2005 MPs enable this monitoring:

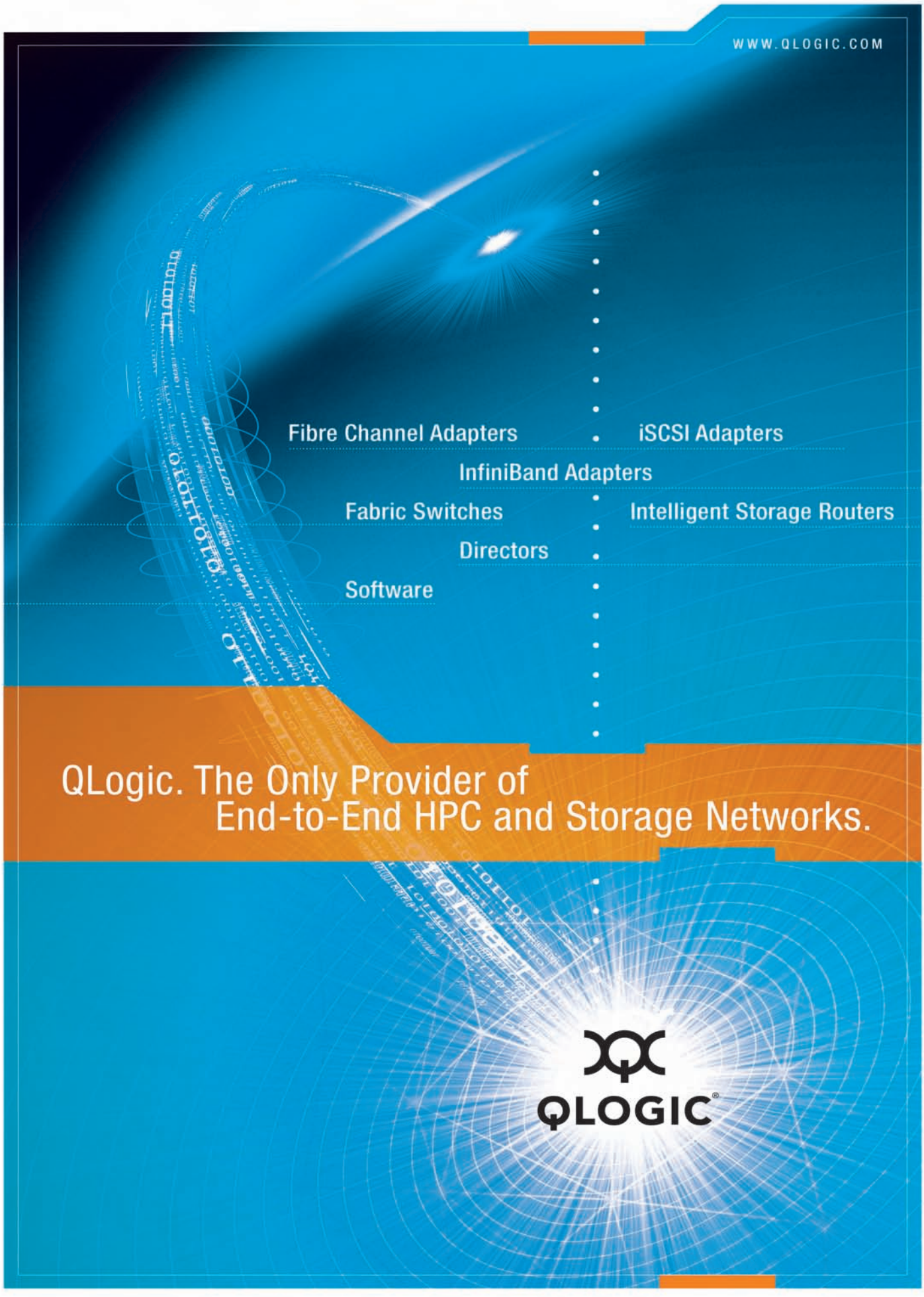
- **Windows Base OS (Server) 2000 and 2003 MP:** Provides parameters and thresholds for monitoring Windows Server–based systems
- **Dell MP:** Enables MOM to detect Dell OpenManage Server Administrator–based hardware alerts for events such as voltage probes and disk failures

Dell IT selected and calibrated the relevant settings and alerts in each MP, then exported them—the Dell implementation of the Windows Base OS MP, for example, included monitoring for problematic event IDs, performance counters, and other key performance indicators such as CPU utilization. ("Server down" alerts, generated when a server suddenly fails catastrophically or becomes unavailable, are provided by the MOM management servers independent of any MP.) These exported MPs represent the standard copy of the monitoring rules, become part of the Definitive Software Library, and are subsequently distributed globally to Dell IT servers without regard to server purpose. Limiting the number of alerts helps support groups effectively manage their environments without the distraction of non-actionable alerts, and MOM consoles enable teams to further instrument monitoring for specific applications or services.

Dell IT also utilizes other MPs, including the SQL Server MP, and plans to introduce MPs for Exchange Server 2003 and the Microsoft Active Directory® directory service in the future.

Consoles

MOM 2005 includes two main consoles: the Administrator console and the Operator console. The Administrator console enables



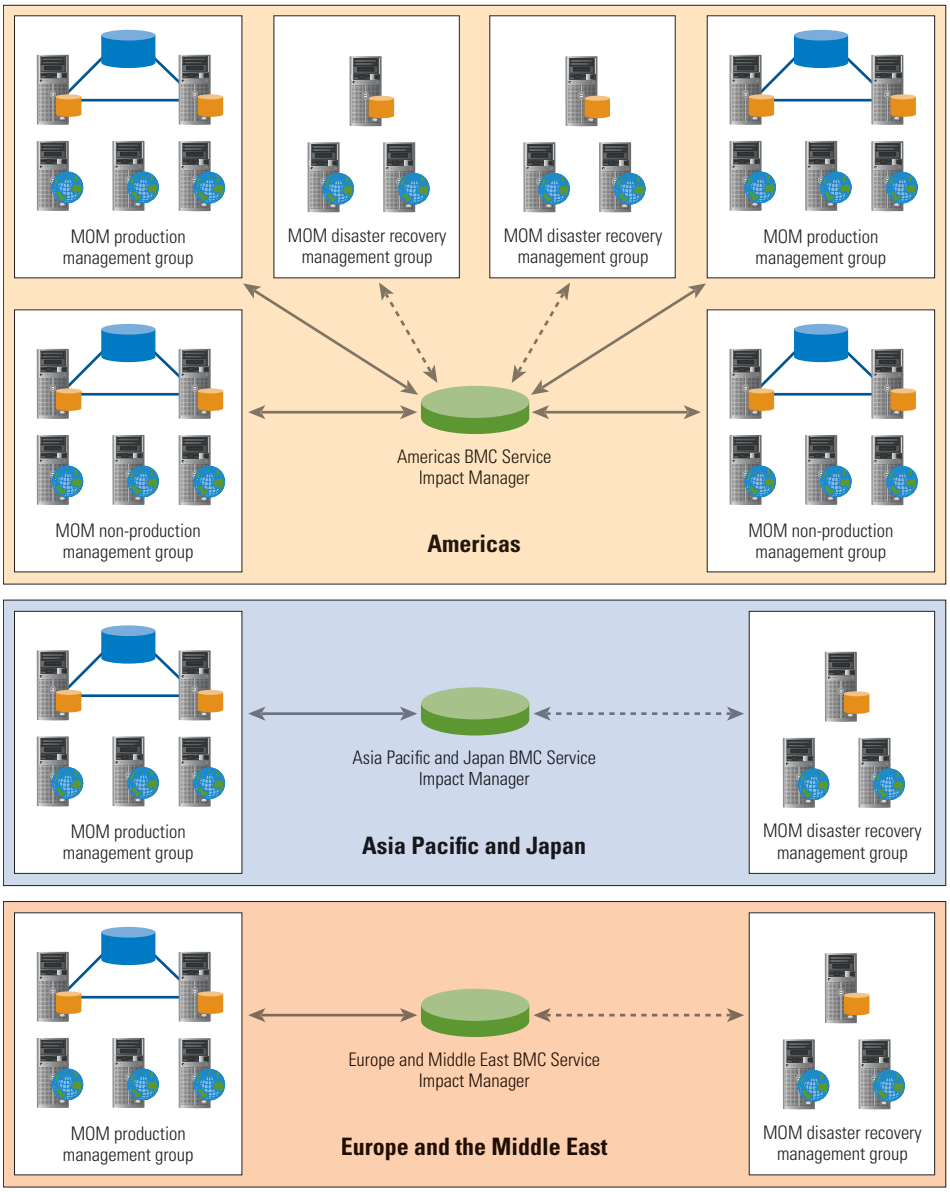


Figure 3. Global Dell IT monitoring infrastructure deployment

support groups to create new monitoring rules for parameters and thresholds, providing the flexibility to create overrides and consolidate rules into rule groups for mass distribution to agents. Support groups can also use this console to create tasks for gathering data on agents and to create custom console views to adapt the Operator console to specific needs.

The Operator console is a Web-based interface that allows support groups to view detailed information on MOM alerts and notifications. Because Dell IT support groups use the BMC Remedy Action Request (AR) System as their primary ticketing console, they use the Operator console mainly to view additional detailed information or troubleshoot problematic systems. Although the Operator console provides fewer features than the Administrator console, its Web-based interface has the advantage of allowing access from anywhere on the Dell IT network.

Supporting infrastructure elements

To provide functionality not provided by MOM 2005 alone, Dell IT also included several other elements in its infrastructure design, including BMC Service Impact Manager (SIM) for event correlation and BMC Remedy AR System for ticketing, BMC Impact Database Gateway for event reporting, custom configuration reporting software, and custom integrated server discovery and asset management software.

BMC Service Impact Manager and Remedy Action Request System

SIM is the Dell IT standard for automatically creating trouble tickets from monitoring applications and sending them to AR System. Neither Microsoft nor BMC provides specific software to allow the integration of MOM 2005 with SIM, but they do provide application programming interfaces (APIs) to support the development of custom software. To help meet its schedule demands, Dell IT purchased an off-the-shelf product connector, the Seamless Technologies Integration of BMC Impact Manager to MOM, and included it in the management group design. This connector, based on the SIM API and MOM Connector Framework (MCF), essentially adds an event response type that forwards the configured event to SIM, which then applies the processing rules and sends the appropriate trouble ticket to AR System.

Two management servers in each management group run the connector in an active/passive configuration. A monitoring system generates alerts if the connector fails, and allows support personnel to manually fail over the connector. Every IT region has its own SIM instance to process events for that region, which are also sent to a global SIM instance in the Americas. Each management group has at least one primary connector and one backup connector to that region's instance.

Processing events in a middleware system such as SIM enables Dell IT to correlate events across monitoring applications. SIM also allows the inclusion of logic in ticket creation, enabling different levels of severities and different types of tickets based on a set of criteria known as the event condition formula. This feature builds on the ability of MOM to detect and generate alerts for errors by providing extensive flexibility in how such errors are translated into trouble tickets and presented to support groups.

BMC Impact Database Gateway

For event reporting, Dell uses BMC Impact Database Gateway, which offloads configured events from SIM into a database and enables powerful reporting and data-mining capabilities. Because Dell IT integrates MOM 2005 with SIM using the Seamless Technologies product connector, reporting events originating from MOM is straightforward. BMC Impact Database Gateway is used to capture events from the global SIM instance and archive them into a global centralized database to provide a single source for relevant global event data. Dell IT configures events requiring reporting to be forwarded by each regional SIM instance to the global instance.

Custom configuration reporting software

A hierarchical MOM 2005 management group design using the included MOM-to-MOM Product Connector can provide a straightforward way to manage an entire environment from a single console, but the scale of the Dell IT infrastructure does not allow such a design. The management groups are therefore unaware of each other, making locating and managing individual agents a daunting prospect.

Dell IT determined that custom software was necessary to help simplify configuration reporting across management groups, and developed a collection of applications and Web interfaces using the MOM Software Development Kit (SDK) and .NET platforms such as C# and ASP.NET for this purpose. Dell IT used the MOM SDK—which provides a collection of classes and methods for accessing MOM 2005 management groups, servers, and databases—to create a single Web portal that could retrieve the configuration for any agent in the global environment regardless of management group.

Custom integrated server discovery and asset management software

Although MOM 2005 includes substantial server discovery capabilities, none are well suited for the scale of the Dell IT environment—an unqualified discovery process could cause problems for wide area network (WAN) links and small, remote sites. Dell IT developed custom integration software that uses the MOM manual installation file, manualmc.txt, to enable a separate application to discover servers present only in the asset management database system that also include the appropriate rule sets according to the class of server. The class of server is determined by values in BMC Remedy Asset Management Application records.

The manualmc.txt file helps provide detailed control over which agents a particular management group monitors, a feature Dell IT used to develop a middleware application to determine the server placement based on asset management database fields, server load, and so on. The application contains logic for geographic placement, server role determination, and load balancing based on records in the Asset Management Application implementation and data obtained using the MOM SDK, and can also remove monitoring for decommissioned servers without requiring support group intervention. In addition, the MOM SDK gathers installation failures and reports them centrally in the same Web portal, providing information such as server name, management group, and failure code along with a reason for the failure.

Integrated global monitoring using MOM 2005

Large-scale enterprise environments often cannot take advantage of off-the-shelf products—and operations that are straightforward to achieve in small environments can require complex, rigorous deployments in large ones. MOM 2005 provides the scalability and features to serve as a solid framework for large-scale enterprise monitoring. It also provides formidable and robust application interfaces, such as the MOM SDK and MCF, which allow it to scale to large server environments using common development platforms.

The Dell IT global monitoring infrastructure demonstrates that these interfaces can allow MOM 2005 to scale effectively to large Windows Server environments. Combined with Dell PowerEdge servers, Dell PowerVault storage, Dell OpenManage Server Administrator, and BMC software, this infrastructure enables Dell IT to effectively monitor its global Windows Server environment.

David Rubio is a systems engineer in the Global Technology and Engineering Services organization in Dell IT. He holds Microsoft Certified Systems Engineer (MCSE) and IT Infrastructure Library (ITIL) Foundation certifications. David has a B.A. in English from the University of New Mexico.

Lon Alonzo is a systems engineer in the Global Technology and Engineering Services organization in Dell IT. He has more than 10 years of IT experience. He holds MCSE and ITIL Foundation certifications.

Pragnesh Rathod is a project manager in the Global Technology and Engineering Services Project Management Office in Dell IT. He holds Project Management Professional and ITIL Foundation certifications. Pragnesh has a B.E. in Industrial and Production Engineering from Gulbarga University and a master's degree in Management Information Systems from Texas Tech University.

Building a Single Image for Dell PowerEdge Servers

Using Altiris Deployment Solution

Using hardware-independent images for OS deployment can free administrators from creating and maintaining separate images for each hardware vendor and server model in their environment, helping significantly simplify server provisioning processes. Altiris® Deployment Solution™ software enables administrators to create and deploy these images on Dell™ PowerEdge™ servers.

BY ERIC SZEWCZYK

Related Categories:

Altiris

Dell PowerEdge servers

Image deployment

Microsoft Windows

System deployment

Systems management

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

Image-based OS deployment processes typically require creating, maintaining, and deploying separate images for each hardware vendor and server model in an IT environment. Using hardware-independent images, however, can eliminate this need, thereby helping simplify server provisioning preparation and deployment processes. Administrators can use Altiris management software to create and deploy these images as part of a comprehensive build process for bare-metal server provisioning—including configuring the BIOS, Dell Remote Access Controller (DRAC), baseboard management controller (BMC), and RAID components; setting up image-based or scripted deployments; remotely deploying server applications and platforms such as Dell OpenManage™ Server Administrator, Microsoft® SQL Server™ 2005, Oracle® Database 10*g*, and VMware® ESX Server; and patching hardware with Dell Update Packages (DUPs).

This article focuses on creating hardware-independent images for Dell PowerEdge servers running Microsoft Windows Server® 2003 operating systems, but includes methodologies that can be applied to the Microsoft Windows® XP and Windows 2000 operating systems as well. One key enabler of hardware-independent images is a common hardware abstraction layer (HAL) for the servers being provisioned. Because Dell PowerEdge servers from the third generation through the current ninth generation share a common HAL based on the Advanced Configuration and Power Interface (ACPI), administrators in data centers using PowerEdge servers can take advantage of the methods described in this article for building and testing these images.

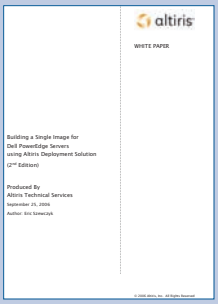
Identifying the benefits of hardware-independent images

Enterprise IT staff typically invest substantial time and effort in system deployment and management, including tasks such as the following:

- Building and maintaining separate images for each hardware vendor, hardware model, and OS for image-based deployments

LEARN MORE ABOUT ALTIRIS DEPLOYMENT SOLUTION

This article is excerpted from the Altiris white paper "Building a Single Image for Dell PowerEdge Servers Using Altiris Deployment Solution," which provides detailed step-by-step instructions, best practices, and links to supporting files such as an image creation checklist, the Microsoft Sysprep utility, a sample Sysprep.inf file, Dell drivers, and the Altiris Deployment Solution AClient. The complete white paper and testing matrix are available at apl-ibase.altiris.com/resources/dell/single_image/WP_Single_Image_for_Dell_Servers_v2.pdf.



- Deploying and patching operating systems, software, and hardware
- Performing regular version-control maintenance—including service packs, security updates, and hot fixes—to keep software up-to-date on both static images and production systems
- Avoiding or managing inconsistent computer builds deployed by different IT staff members (including elements such as driver versions, configuration methods, and overall look-and-feel)

Maintaining and deploying hardware-independent images for Dell PowerEdge servers helps simplify many of these types of tasks by offering the following potential benefits:

- **Efficient image management:** IT staff only need to maintain a single image for all PowerEdge servers, rather than separate images for each model, which helps significantly reduce the time required for bare-metal OS deployments.
- **Flexible automated deployment:** In addition to using Altiris Deployment Solution, administrators can deploy self-extracting images through a bootable CD or DVD, network share, USB drive, storage area network, and so on.[1]
- **Simplified updates:** Administrators can use Altiris Patch Management Solution™ software to identify software and hardware vulnerabilities and remotely deploy current service packs, security updates, hot fixes, and DUPs once they have installed the Altiris agent and moved the system into production.
- **Robust version control and licensing:** Rather than preinstalling applications in a static image and increasing the frequency of image maintenance to keep these applications up-to-date, administrators can repackage applications with Altiris RapidInstall™ software (included with Altiris Deployment Solution) or Wise Package Studio (acquired by Altiris), which enables them to maintain control over application updates and licenses separately from the OS image. They can then deploy the applications with Altiris Deployment Solution or Altiris Software Delivery Solution™ software using push methodologies to help reduce the time spent on image updates.
- **Consistent builds:** Using the same image across different PowerEdge servers enables systems to maintain the same foundational drivers, configurations, and look-and-feel.

Images provide one important method of OS deployment, but administrators also have another method available: scripted installations. One method may be more suitable than the other based on the network environment, which is in turn dictated by complex IT policies, diverse network architectures, IT staff preferences, and so on. Altiris software helps administrators perform both types of deployments. Figures 1 and 2 summarize the advantages and disadvantages of each method.

Building hardware-independent images

Altiris Deployment Solution enables administrators to easily capture OS images and deploy them to multiple systems. It also serves as the environment for add-on products such as Altiris Deployment Solution for Dell Servers, which helps administrators perform bare-metal provisioning of Dell PowerEdge servers as well as post-deployment

Advantages of image-based deployment	Disadvantages of image-based deployment
• *Minimal deployment time:* A 2 GB 16-bit DOS image might take only 6–8 minutes to deploy, and a 2 GB 32-bit Windows Preinstallation Environment (WinPE) or Linux® OS image might take only 2–3 minutes to deploy. • *Efficient distribution:* Images can be multicasted to multiple systems simultaneously, which helps reduce network bandwidth and bottleneck constraints based on the broadcast method. • *Single image:* A single image can be created for different Dell PowerEdge models, eliminating the need to create and maintain separate images for each OS and model. • *Consistent builds:* Administrators can maintain OS build consistency without intervention or post-deployment scripts.	• *Inflexible updating:* After the image has been captured, the software and drivers contained in the image become outdated over time and require updates. Typically, administrators must either regularly deploy, update, and recapture the master image, or load management agents in the image and deploy updates later with policy-based distributions.

Figure 1. Advantages and disadvantages of image-based OS deployments

[1] For more information, see the "Option 2 – Deploying the Image Independent of Deployment Solution" section in the complete Altiris white paper at apl-ibase.altiris.com/resources/dell/single_image/WP_Single_Image_for_Dell_Servers_v2.pdf.

tasks such as software installation, OS and hardware patching, and inventories. This section outlines some basic tips for building hardware-independent images.

Identify HAL compatibilities. When deploying hardware-independent images, the master and destination server HALs must be compatible. For example, an image captured from a non-ACPI-compliant system and deployed to an ACPI-compliant system typically fails or becomes caught in a continuous boot cycle. Fortunately, most PowerEdge servers—those from the third generation through the current ninth generation—have the same HAL, so incompatibility typically is not a problem when creating and deploying images on Dell servers.

Identify a reference server for image creation. To help simplify image creation and deployment to target hardware, administrators should select a standard PowerEdge server model as a reference server from which to create the image. This server should not include any superfluous hardware or configurations—for example, it should use SCSI rather than RAID, and DRACs and other add-on cards should be removed or disabled. If not using the Altiris Deployment Solution for Dell Servers add-on to provision the hardware, or if using a PowerEdge model not supported by the Dell OpenManage Deployment Toolkit and therefore not supported by this add-on, administrators can initially provision the hardware and deploy the OS using Dell OpenManage Server Assistant, which is included on a CD with PowerEdge servers and available for download from support.dell.com.

Advantages of scripted deployment	Disadvantages of scripted deployment
• *Flexible updating:* Service packs, hot fixes, and driver updates can be added to the installation process to help keep OS deployments up-to-date. • *Simple customization:* Administrators can easily customize components during installation.	• *Lengthy deployment time:* OS deployments can take 2–4 hours depending on the number and size of the updates. • *Inefficient distribution:* Installations cannot be multicasted to multiple systems simultaneously. • *Tedious update maintenance:* The process of continually maintaining the latest updates and slipstreaming them into the OS deployment can be monotonous. (One alternative is to use patch management software such as Altiris Patch Management Solution for Windows, which allows the management agent to identify OS vulnerabilities after deployment, and then push out the updates as part of a policy-based distribution.) • *Time-consuming preparation:* Administrators must perform initial preparatory tasks on the deployment server, such as copying media files and drivers and editing answer files, before they can use predefined scripted OS jobs.

Figure 2. Advantages and disadvantages of scripted OS deployments

Many hardware components are embedded on the motherboard and can be disabled either from the system BIOS or by opening the case and physically removing the modules. (Administrators should consult the hardware manual and warranty before attempting to physically remove modules themselves.) For example, a server with an embedded PowerEdge Expandable RAID Controller (PERC) can be disabled in the system BIOS, which then defaults to using SCSI. Disabling a PERC allows administrators to build an image from one SCSI drive with a small OS partition, rather than three drives with a typical RAID-5 configuration, thereby simplifying the image build to help increase the success rate when mass deploying.

Use the minimum available hard disk or partition size. Typically, the source hard drive or partition used to create an image should not be larger than the target hard drive or partition; for example, administrators should not create an image from an 80 GB drive or partition and deploy it to a 40 GB drive or partition. One way to avoid potential problems is create a small partition for image creation. For example, a 6–8 GB partition typically allows space for the OS, the .\i386 directory copied from the Windows media CD, and any service packs, security updates, and hot fixes—while still providing some extra space if needed.

Select a target server for image testing. To verify that an image has been configured properly, administrators should test the image on a different PowerEdge model with different chipsets and hardware than the reference server. If no such server is available, administrators can also test the image on the reference server by re-enabling all the hardware components and configurations that had been disabled, if applicable. This method is not as reliable as using an entirely different model, however, because even though the Sysprep utility should remove the current hardware devices installed on the reference server, the drivers are still present in the system directories and could be reinstalled when the system boots up again.

Obtain the appropriate version of Microsoft Sysprep. There are currently three versions of Sysprep available from Microsoft: version 1.0, included with the Windows 2000 product CD; version 1.1, the current version for Windows 2000; and Version 2.0, the current version for Windows Server 2003 and Windows XP.[2]

Use a volume license key. Administrators should use volume license keys and media for the OS that the hardware-independent image will be built from, which avoids Windows activation on the target systems.

Identify target system hardware. To successfully build a hardware-independent image, administrators should identify the different hardware devices that exist in the environment, such as DRACs, PERCs, modems, network interface cards, and so on.

[2] Version 2.0 is included in the supporting files for the complete Altiris white paper, available at apl-ibase.altiris.com/resources/dell/single_image/WP_Single_Image_for_Dell_Servers_v2.pdf. For more information about Sysprep, see the "Introduction to Microsoft Sysprep" section in the complete white paper.

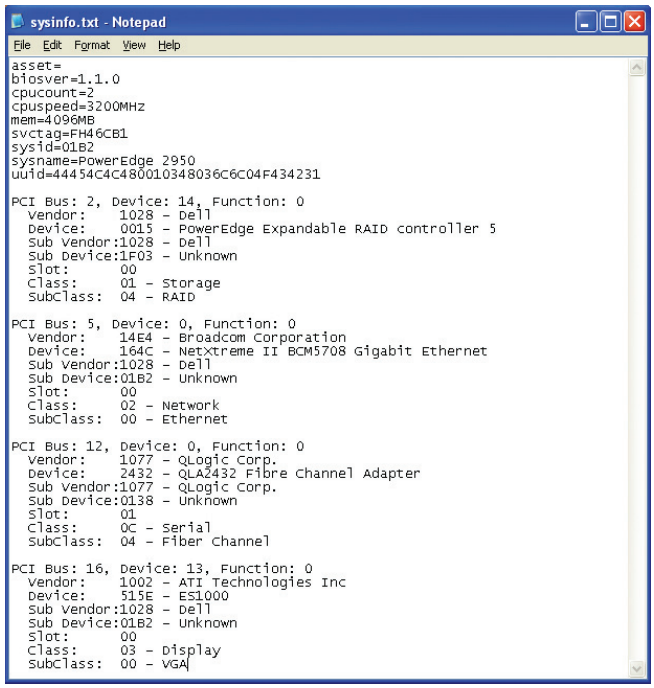


Figure 3. Altiris Deployment Solution for Dell Servers hardware inventory output for a Dell PowerEdge 2950 server

Altiris Inventory Solution® software, Altiris Inventory Solution for Servers (an upcoming release), Altiris Real-Time System Manager Solution™ software, and even the Altiris Deployment Solution for Dell Servers add-on can identify the hardware that exists in an environment. For example, Altiris Deployment Solution can capture commonly requested inventory through the Altiris AClient agent and display it through the Altiris Deployment Solution console. The Altiris Deployment Solution for Dell Servers add-on can also use the Dell OpenManage Deployment Toolkit (DTK) to capture and output detailed hardware inventory to a text file (see Figure 3), and can run in a pre-OS environment without an agent.

After gathering this information, administrators can download the appropriate vendor-specific or third-party Plug and Play (PnP) drivers for each hardware device, chipset, and so on, so that they can properly gather and bundle these drivers into the single image for remote installation during the PnP detection process. Although some or most of these drivers may be native to the OS itself, administrators typically should download the drivers from the hardware vendor itself—in the case of PowerEdge servers, from support.dell.com.[3]

Identify legacy devices. Legacy devices are those that do not support PnP and may require manual installation and configuration after image deployment. Identifying these devices beforehand helps save time and effort later, and may help identify an automated work-around to avoid manual intervention—for example, because Altiris Deployment Solution can deploy any type of script in a pre-OS or post-OS environment, administrators may be able to set up an unattended script that Altiris Deployment Solution can then deploy without manual intervention.

Properly build the SysprepMassStorage section of Sysprep.inf. One of the most critical (and tedious) parts of creating a hardware-independent image is properly building the `SysprepMassStorage` section of the Sysprep.inf file, which identifies elements such as the PnP PCI vendor IDs extracted from the vendor-specific .inf files, the location of the drivers contained in the image, and so on, so that they can be properly identified and installed during the post-deployment mini-setup process. Building this section incorrectly may cause errors during the pre-deployment Sysprep process if the OS cannot properly identify and load the appropriate mass storage controller drivers.[4]

For an example Sysprep.inf file, see the supplemental online section of this article at www.dell.com/powersolutions.

Simplifying image-based OS deployment

Using hardware-independent images helps dramatically simplify image-based server provisioning—including provisioning of Dell PowerEdge servers. Altiris Deployment Solution along with other Altiris software tools can provide a flexible way for administrators to create and deploy these images. Implementing Altiris software in conjunction with the methods outlined in this article enables an efficient, robust deployment process for enterprise data centers.

Eric Szewczyk is a technical strategist on the Dell Alliance at Altiris. He manages the Dell IT relationship in addition to training Dell systems engineers to use Altiris management software. Eric has a B.A. from the University of Central Oklahoma and is an Altiris Certified Engineer (ACE).

FOR MORE INFORMATION

Szewczyk, Eric. "Building a Single Image for Dell PowerEdge Servers Using Altiris Deployment Solution." Altiris Technical Services, September 25, 2006. apl-ibase.altiris.com/resources/dell/single_image/WP_Single_Image_for_Dell_Servers_v2.pdf

Supporting files for complete Altiris white paper:
apl-ibase.altiris.com/resources/dell/single_image/onedellimage.zip

Altiris and Dell server deployment:
www.altiris.com/delldeploy

Altiris and the Dell Alliance:
www.altiris.com/dell

Altiris products:
www.altiris.com/products

[3] For more information, see the "Gathering Dell Drivers for Inclusion in the Hardware Independent Image" section in the complete Altiris white paper at apl-ibase.altiris.com/resources/dell/single_image/WP_Single_Image_for_Dell_Servers_v2.pdf.

[4] A sample Sysprep.inf file and Dell drivers are included with the supporting files for the complete Altiris white paper, available at apl-ibase.altiris.com/resources/dell/single_image/WP_Single_Image_for_Dell_Servers_v2.pdf. For more information and step-by-step instructions on properly creating this section, see the "Building the [SysprepMassStorage] Section of the SYSPREP.INF File" section in the complete white paper.

Simplifying Blade Server Management

with Altiris Deployment Solution Rip and Replace

Altiris has extended its Altiris® Deployment Solution™ software to offer automated rip-and-replace functionality for Dell™ PowerEdge™ blade servers. This article describes how to configure and use this feature to help simplify blade server deployment.

BY ERIC SZEWCZYK AND CHAD FENNER

Related Categories:

Altiris

Blade servers

Dell PowerEdge blade servers

System deployment

Systems management

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

Many enterprises today struggle with the rising costs of server management. Blade servers can provide several advantages in enterprise data centers, including allowing administrators to use fewer cables and, after the chassis is racked, carry out faster physical deployments than they can with 1U rack servers. Dell has teamed with Altiris to provide blade servers with truly simplified management, helping reduce data center complexity by easing blade deployment and relocation within a chassis.

Altiris server management software is designed to simplify common tasks such as server deployment, software and firmware updating, inventory, monitoring, security auditing, and asset management, and uses a single console for easy centralized management. Because Altiris software is designed to support Dell blade, rack, and tower servers equally, administrators can manage blade servers using the same console, agents, and policies they use for rack and tower servers. Using Altiris software in enterprise data centers can dramatically reduce initial server deployment time,[1] helping simplify data center operations and reduce IT management costs. And support for Dell blade servers does not incur additional licensing costs—administrators can add Dell support for existing licensed installations of Altiris Deployment Solution by installing the free Altiris Deployment Solution for Dell Servers add-on module.[2]

One key feature of blade servers is the ability to associate a specific server with a specific location in a chassis. To take advantage of this characteristic, Altiris has extended its standard server management software to include additional automation functions for blade servers. Chief among these is the Altiris Deployment Solution rip-and-replace feature.

Configuring Altiris Deployment Solution rip and replace

Rip and replace refers to automating blade server deployment so that administrators can quickly and easily deploy a replacement blade with the same configuration as the previous blade. This deployment can be as simple as removing the old blade from the chassis, then inserting the new one in the same slot and powering it up.

[1] For more information, see "Time-Savings Validation for Dell Server Deployment with Altiris Deployment Solution," by Todd Mitchell and Landon Hale, in *Dell Power Solutions,* August 2005, www.dell.com/downloads/global/power/ps3q05-20050221-Altiris.pdf.

[2] This module is available for download at www.altiris.com/eval/dell.

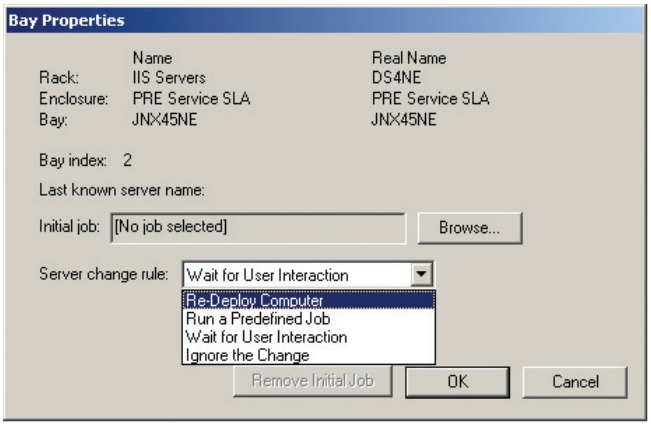


Figure 1. Server change rules in the Bay Properties window of the Altiris console

Administrators can configure Altiris Deployment Solution to detect Dell service tags, Media Access Control (MAC) addresses, asset tags, or universally unique identifiers (UUIDs) during boot and use that information to perform specialized bare-metal server builds. Because servers may include multiple network adapters, Dell service tags are typically recommended as the primary lookup key; administrators can set this option in the Altiris console by selecting Tools > Options > Global > Primary lookup key.

In addition to providing deployment functionality, Altiris Deployment Solution enables powerful post-deployment management, including software delivery, inventory, hardware reconfiguration, and rapid server repurposing—the last of which is a key function of the rip-and-replace feature.

Server change rules

Administrators can configure the rip-and-replace feature in the Altiris console using *server change rules,* which designate the action the software should take if it detects a server change in a particular chassis slot. Chassis icons are displayed in the Physical Devices tree of the Computers pane in the Altiris console; administrators can assign change rules either in the Bay Properties window for the chassis (see Figure 1) or in the properties window for an individual chassis slot. They can choose from four rules:

- **Re-Deploy Computer:** Altiris Deployment Solution automatically deploys the new blade using the last sequence of deployment and configuration jobs in the previous blade's job history,[3] without requiring administrators to begin the build process through the Altiris console. This setting enables administrators to quickly and easily replace a failed blade.
- **Run a Predefined Job:** Altiris Deployment Solution automatically runs a predefined job on the new blade. This job can employ different levels of deployment, including various combinations of hardware configuration, OS deployment, and application installation. This type of automation can reduce the time administrators spend on basic installation functions from hours to minutes.[4]
- **Wait for User Interaction:** This is the default setting, in which Altiris Deployment Solution waits for administrators to manually perform deployment tasks. For blades previously deployed elsewhere, Altiris Deployment Solution does not automatically associate its history or configuration parameters with its new slot, and the Altiris agent on the blade waits for further instructions. An icon in the Altiris console indicates that the blade is waiting until an administrator drags and drops the first job on the blade.
- **Ignore the Change:** Altiris Deployment Solution does not run any jobs or other automated rip-and-replace functions, but does associate the history and configuration parameters of blades previously deployed in another slot with the new slot—a useful option when administrators want to move blades to different slots while maintaining their existing configuration. If the blade has not been deployed elsewhere, then the typical default deployment mechanisms are available, such as initial deployment through the Preboot Execution Environment (PXE) menu. This type of globally defined deployment process operates separately from rip-and-replace functions.

The Altiris console supports two views for working with blade servers: an administrator-defined hierarchy showing server groups, which appears in the All Computers tree in the Computers pane; and a physical hierarchy showing racks, chassis, and slots, which appears in the Physical Devices tree in the Computers pane. Administrators can use the server groups to manage blades just like rack and tower servers, or use the physical hierarchy view to drag and drop jobs onto server blades installed in a particular rack, chassis, or slot.

To assign server change rules, administrators can first create a chassis view in the Altiris console by right-clicking on the Physical Devices tree in the Computers pane and selecting "New Virtual Bay." In the Create Virtual Bays window, they can enter rack and enclosure names, select the enclosure type from the drop-down menu, and set the server change rule for the entire chassis (see Figure 2). After creating the chassis view, they can also set rules for individual slots.

[3] Altiris Deployment Solution maintains a detailed history of management functions executed on a managed computer, including configuration changes, imaging events, custom scripts, software delivery tasks, and so on. When the Re-Deploy Computer rule is used on a newly inserted blade, Altiris Deployment Solution executes the tasks in the slot's history starting with the last Distributing a Disk Image task or Scripted OS Install task, or from the last Run Script task that includes the `rem deployment start` command.

[4] For more information, see "Time-Savings Validation for Dell Server Deployment with Altiris Deployment Solution," by Todd Mitchell and Landon Hale, in *Dell Power Solutions,* August 2005, www.dell.com/downloads/global/power/ps3q05-20050221-Altiris.pdf.

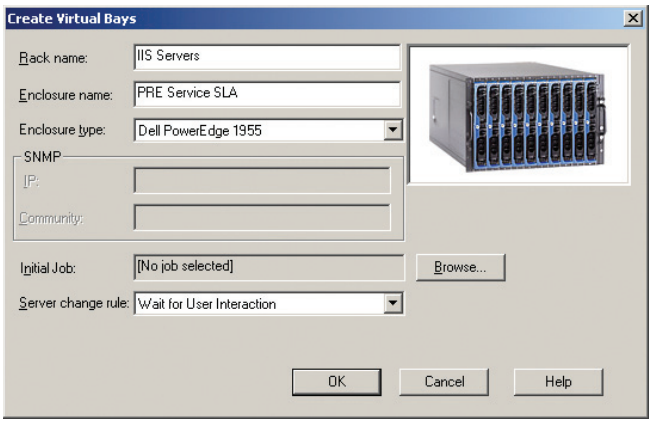


Figure 2. Create Virtual Bays window in the Altiris console

Performing rip-and-replace operations

Performing rip-and-replace operations by replacing an existing blade with a new one involves the following steps:

1. Administrators remove a blade server from a chassis slot, then install the new blade in that slot.
2. Administrators power up the new blade manually or by using remote power control through the Dell Remote Access Controller (DRAC) or baseboard management controller Intelligent Platform Management Interface (IPMI) interfaces available as part of the Altiris Deployment Solution for Dell Servers add-on module.
3. The blade network interface card that is set to network boot connects to the Altiris PXE server.[5]
4. The Altiris PXE server checks the blade's MAC address to determine whether it recognizes the blade as an existing Altiris managed device. If the PXE server recognizes the blade, then it directs the blade to load a managed image. If the PXE server does not recognize the blade, then it directs the blade to load the Initial Deployment PXE boot image, and the procedure continues with the following steps.
5. The blade boots from the Initial Deployment image, and the Altiris agent within that image contacts Altiris Deployment Server.
6. Altiris Deployment Server confirms that the blade is new using the primary lookup setting from the Altiris console, then executes the server change rule for the appropriate chassis slot. If administrators are using the Re-Deploy Computer rule, the deployment process is automated and requires no manual intervention.

Avoiding unwanted deployments

Altiris Deployment Solution is designed to avoid mistaken blade re-imaging or redeployment with the rip-and-replace feature. For example, to help ensure that the software does not perform unwanted jobs or other deployment tasks automatically, the default change rule for each chassis slot is Wait for User Interaction. In addition, the Altiris Initial Deployment feature is disabled by default for new servers, which helps prevent deployment events from executing on managed or unmanaged servers without explicit permission. Administrators enable this feature by right-clicking on the Initial Deployment job, selecting Properties > Advanced, and clearing the Servers check box.

Simplifying blade server management

The rip-and-replace feature for Dell PowerEdge blade servers is a standard part of Altiris Deployment Solution for Dell Servers, and can provide significant advantages for both new and existing users of these servers. Using this software to manage blades does not require special licensing beyond the standard Altiris Deployment Solution per-server licenses. Once implemented, Altiris Deployment Solution can automatically detect Dell blade servers as they are added to the environment, helping provide a simplified, cost-effective way to manage these servers.

Eric Szewczyk is a technical strategist on the Dell Alliance at Altiris. He manages the Dell IT relationship in addition to training Dell systems engineers to use Altiris management software. Eric has a B.A. from the University of Central Oklahoma and is an Altiris Certified Engineer (ACE).

Chad Fenner is the product marketing manager for blade servers at Dell. He has a bachelor's degree from Trinity University in San Antonio, Texas.

FOR MORE INFORMATION

Altiris and Dell:
www.altiris.com/dell
www.dell.com/altiris

Altiris Deployment Solution for Dell Servers:
www.altiris.com/delldeploy

Altiris Deployment Solution for Dell blade servers:
www.altiris.com/upload/ds_dsfornewdellservers_6.pdf

Altiris Deployment Solution documentation:
www.altiris.com/support/documentation

Altiris Deployment Solution video:
apl-ibase.altiris.com/resources/dell/demo/Deployment10min.wmv

Dell Inc. "Increasing Efficiencies for One of the World's Most Efficient Companies." December 2005. www.altiris.com/upload/dell_it_success_story.pdf

Mitchell, Todd, and Landon Hale. "Time-Savings Validation for Dell Server Deployment with Altiris Deployment Solution." *Dell Power Solutions,* August 2005. www.dell.com/downloads/global/power/ps3q05-20050221-Altiris.pdf

[5] For more information about the PXE boot process, see the Altiris Deployment Solution documentation at www.altiris.com/support/documentation.

Enabling Server Consolidation with

64-Bit Citrix Presentation Server

Deploying 64-bit technology can allow organizations to significantly reduce the number of servers required to support their user base by enabling increased density and resource utilization. This article discusses the advantages of running the 64-bit version of Citrix Presentation Server™ software for Microsoft® Windows Server® 2003 x64 Editions on Dell™ PowerEdge™ servers.

BY JOHN D'AGATI AND MICHAEL EAST

Related Categories:

Citrix Systems

Dell PowerEdge servers

Microsoft Windows

Microsoft Windows Server 2003 x64 Editions

Server consolidation

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

The limitations of 32-bit technology, particularly its restrictions on the amount of memory that can be directly addressed, can prevent software from taking full advantage of powerful multi-core processors. The introduction of 64-bit hardware and software has helped reduce or eliminate these limits, enabling significant scalability and performance benefits. The 64-bit version of Citrix Presentation Server 4 is designed to take advantage of these benefits on multi-core servers running Microsoft Windows Server 2003 x64 Editions.

The capabilities of the 64-bit version of Citrix Presentation Server 4 along with Windows Server 2003 x64 Editions can enable a compelling return on investment[1] while helping reduce total cost of ownership through data center server consolidation. Enterprises can now take advantage of servers with four or more processor cores, such as Dell PowerEdge 2950 servers, in 64-bit Citrix Presentation Server 4 environments.

Advantages of 64-bit technology

In 32-bit Microsoft Windows® operating systems, memory is limited to the 32-bit address space, thus limiting the amount of virtual memory that can be directly addressed to 4 GB. This 4 GB is divided into two equal parts: 2 GB for application processes and 2 GB for the OS, which is used for the kernel memory, system cache, and drivers.

[1] For more information, see "Reduce IT Costs Using Citrix Presentation Server 4," by Citrix Systems, October 2005, www.citrix.com/site/resources/dynamic/salesdocs/PresentationServer4.0x64BusinessCaseWhitepaper01DEC2005.pdf.

	32-bit	64-bit
Total physical memory	64 GB	1 TB
Total virtual memory (based on a single process)	4 GB	16 TB
Virtual paged pool	470 MB	128 GB
Virtual non-paged pool	256 MB	128 GB
Virtual page table entry	660 MB to 900 MB	128 GB

Figure 1. Physical and virtual memory limits for 32- and 64-bit versions of Windows Server 2003 Enterprise Edition

	Server	Client
Model	Dell PowerEdge 2950	Dell PowerEdge 1750
Processors	Two dual-core Intel® Xeon® 5160 processors at 3.0 GHz	One Intel Xeon processor 80532K at 2.8 GHz
Cache	Four 2 MB L2 caches (one per core)	One 512 KB L2 cache
Frontside bus	1,333 MHz	533 MHz
Memory	16 GB	1 GB
Hard drives	• Two Serial Attached SCSI (SAS) 73 GB, 15,000 rpm drives in a RAID-1 configuration • One SAS 73 GB, 15,000 rpm drive for the page file	One Ultra320 SCSI 18 GB, 15,000 rpm drive
RAID controller	PowerEdge Expandable RAID Controller (PERC) 5/i	N/A
Network interface	Broadcom NetXtreme II Gigabit* Ethernet adapter	Broadcom NetXtreme Gigabit Ethernet adapter

*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 2. Hardware configuration used in the test environment

	32-bit environment	64-bit environment
OS	Microsoft Windows Server 2003 Enterprise Edition with Service Pack 1 (SP1)	Microsoft Windows Server 2003 Enterprise x64 Edition with SP1
Citrix software	Citrix Presentation Server 4 (32-bit)	Citrix Presentation Server 4 (64-bit)
Applications	Microsoft Office Professional Edition 2003 suite (Excel, Access, and PowerPoint)	

Figure 3. Software used in the test environment

A key advantage of using the 64-bit version of Citrix Presentation Server 4 with each of the Windows Server 2003 x64 Editions is the increased virtual address space compared to the 32-bit versions: the Windows Server 2003 x64 Editions can address 16 TB of virtual memory, divided into 8 TB for application processes and 8 TB for the OS (see Figure 1). This increase in available user and kernel memory enables Citrix Presentation Server to be highly scalable and avoid the limitations of 32-bit architectures.

Performance testing of 64-bit Citrix Presentation Server 4

In September 2006, Citrix engineers used Citrix® ICAMark 3.0 to test the user scalability of the 32- and 64-bit versions of Citrix Presentation Server 4. ICAMark is an internal benchmarking tool based on the Citrix Server Test Kit that tracks the number of client sessions that can be connected to a server running Citrix Presentation Server while still maintaining acceptable performance. It measures the user capacity based on the session response time and calculates a score to help determine the server scalability.

An increase in available user and kernel memory enables Citrix Presentation Server to be highly scalable and avoid the limitations of 32-bit architectures.

The test simulates users performing various actions using Microsoft Excel 2003, Microsoft Access 2003, and Microsoft PowerPoint® 2003 software. In the tests described in this article, all applications run with ICAMark 3.0 were 32-bit.[2] All tests were run until the resource thresholds were met or the end-user experience degraded below a defined threshold. (It is important to remember that server scalability is not always determined by hard system bottlenecks; application responsiveness and user experience should be regarded as the limiting factors.)

Figures 2 and 3 summarize the hardware configuration and software used in the test environment.

Test results

The increase in available kernel memory in 64-bit environments enabled the 64-bit version of Citrix Presentation Server to use multi-core processing power much more effectively than it could in 32-bit environments: in the Citrix tests, the PowerEdge 2950 server running the 32-bit versions of Windows Server 2003 and Citrix Presentation Server 4 supported 211 users, while the PowerEdge 2950 server running the 64-bit versions of Windows Server 2003 and Citrix Presentation Server 4 supported 385 users—an increase of 82 percent.

[2] Windows Server 2003 x64 Editions are designed to execute 32-bit applications by running them inside the Windows on Windows 64 (WOW64) execution layer.

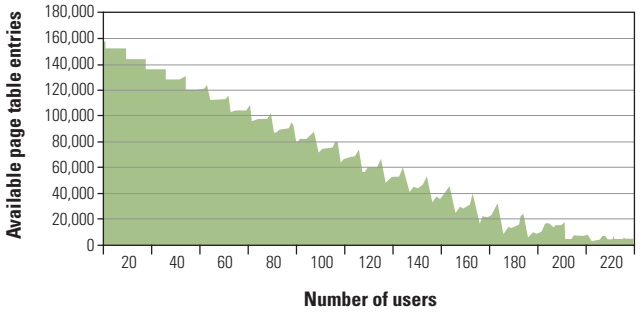


Figure 4. Available page table entries during a test of the 32-bit version of Citrix Presentation Server 4

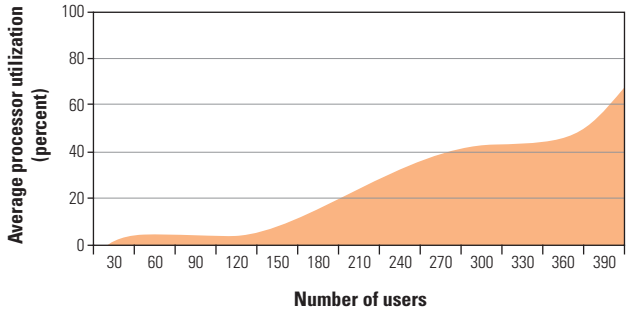


Figure 5. Average processor utilization during a test of the 64-bit version of Citrix Presentation Server 4

Figures 4 and 5 illustrate the effect kernel memory can have on performance. Figure 4 shows the available page table entries for the 32-bit version of Citrix Presentation Server 4. At 211 users, only 1,361 page table entries were available, which is below the recommended value of at least 3,000—and at this point, Citrix Presentation Server began experiencing session launch failures, which can be attributed to a lack of kernel memory resources, and particularly a lack of available page table entries. The bottleneck in this system was caused by the memory management limitations in the 32-bit Windows Server 2003 OS.

The 64-bit version of Citrix Presentation Server, meanwhile, was constrained not by available page table entries but by processing power. Figure 5 shows the average processor utilization for the 64-bit version of Citrix Presentation Server. The system maintained acceptable performance until the number of users reached 385, which was when the large number of users simultaneously requesting resources caused the user experience to degrade because of high contention for resources and increased context switching. At that point, the average processor utilization was consistently above 50 percent, with prolonged spikes at 100 percent, while the other server subsystems—including available kernel memory and page table entries—remained within acceptable limits.

Scalability and cost-effectiveness with 64-bit Citrix Presentation Server 4

One of the key advantages of scaling up by deploying 64-bit technology with multi-core servers is the reduced number of servers required to support end users—for example, the test results indicate that a data center supporting 10,000 users in an environment identical to the one used in the Citrix benchmark tests would need to purchase and maintain 48 servers to run the 32-bit versions of Citrix Presentation Server 4 and Windows Server 2003, but only 26 servers to run the 64-bit versions.

While the kernel memory restrictions of 32-bit architectures severely limit the ability of Citrix Presentation Server to take advantage of multi-core processing power, the 64-bit version of Citrix Presentation Server 4 and Windows Server 2003 x64 Editions can overcome these restrictions, allowing significantly increased scalability and user density. Because utilizing fewer servers typically requires less space, power, and supporting infrastructure; fewer administrators; less time to manage and deploy hot fixes, service packs, and other updates; and fewer software licenses than would otherwise be necessary, deploying 64-bit technology can also ultimately help reduce deployment, operations, and maintenance costs in enterprise data centers.

One of the key advantages of scaling up by deploying 64-bit technology with multi-core servers is the reduced number of servers required to support end users.

John D'Agati is a senior test engineer in the Virtualization Systems Group at Citrix. John is currently the lead for the Citrix Presentation Server scalability and performance team. He has a B.S. in Computer Information Systems from Florida Atlantic University.

Michael East is a senior manager of product development at Citrix, and is responsible for Citrix Presentation Server scalability, performance, and resiliency. He has a B.S. in Computer Engineering from the Georgia Institute of Technology and an M.B.A. from the University of Miami.

FOR MORE INFORMATION

Citrix Presentation Server:
www.citrix.com/English/ps2/products/product.asp?contentID=186

Microsoft Windows Server 2003 x64 Editions:
www.microsoft.com/windowsserver2003/64bit/x64/default.mspx

Customizing Management of HPC Clusters

High-performance computing (HPC) clusters are becoming increasingly widespread as many organizations' large-scale computing needs grow, but each type of cluster configuration comes with its own hardware management requirements. This article discusses best practices for customizing management software and hardware for HPC clusters based on Dell™ PowerEdge™ servers.

BY ARUN RAJAN, TONG LIU, YUNG-CHIN FANG, GARIMA KOCHHAR, AND RON PEPPER

Related Categories:

Cluster management

Clustering

Dell OpenManage

Dell PowerEdge servers

High-performance computing (HPC)

Security

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

High-performance computing (HPC) clusters consist of complex hardware, firmware, and software. In Dell HPC clusters, the hardware typically includes the latest generation of dual-core Intel® Xeon® processors, the latest chipsets, frontside buses with high data transfer rates, the latest memory architectures such as quad-channel fully buffered double data rate 2 (DDR2) memory, and high-performance PCI Express buses. The firmware comprises the Intelligent Platform Management Interface (IPMI), System Management BIOS (SMBIOS), and other industry-standard specifications. The software includes instrumentation services, Advanced Configuration and Power Interface (ACPI) and OpenIPMI modules integrated into the Linux® OS extended kernel, and the latest Dell OpenManage™ suite supporting hardware management from the circuit level to cluster-level out-of-band management.

In addition to these elements, the Platform Open Cluster Stack (OCS) cluster computing software stack (formerly called Platform Rocks) supports cluster software deployment, resource management, monitoring of the cluster-level performance, cluster job state, and resource utilization. This article addresses frequently used HPC cluster management components and best practices for large-scale deployments of HPC clusters.

Using various consoles for cluster management

Clusters can be managed by two types of consoles: one-to-many cluster-level monitoring and management consoles and one-to-one command-line interface (CLI) remote consoles. One-to-many consoles include Dell OpenManage IT Assistant, typically used to monitor and manage cluster hardware health; the open source

Ganglia console, used to monitor cluster-level performance; and the open source Cluster Monitoring (Clumon) console, used to monitor cluster job state and resource utilization.

One-to-one remote consoles are typically used to submit jobs, develop code, and debug. These consoles connect to the baseboard management controllers (BMCs), remote access controllers (RACs), and serial port consoles:

- **BMCs:** Support text-mode console redirection using Serial Over LAN (SOL) Proxy Telnet sessions. The BMC Management Utility (BMU) contains an IPMI shell CLI for communication with the BMC from a remote node, and the SOL Proxy utility for multiple Telnet sessions.
- **RACs:** Provide a dedicated network port and support full-graphics-mode console redirection; these are typically add-on cards. The RAC CLI is called racadm. HPC clusters equipped with Gigabit Ethernet[1] configurations and running Ethernet-based communication-intensive applications can benefit from the RAC dedicated network interface card (NIC), because management traffic can be redirected to this NIC.
- **Serial port consoles:** Support text-mode console redirection using a terminal emulator. Physical serial ports with proper firmware and host OS configuration also support the native IPMI command set.

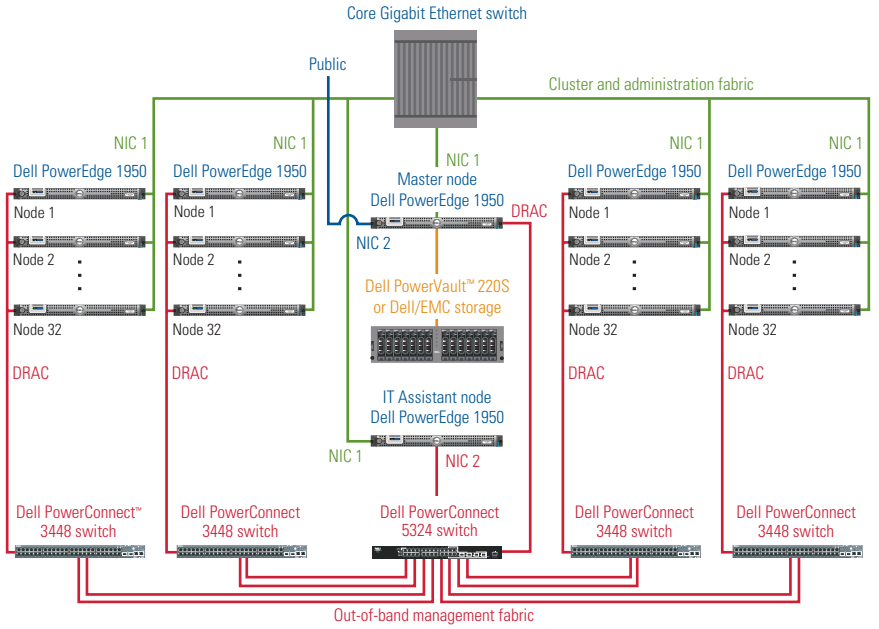


Figure 1. Dell HPC cluster architecture

All three types of one-to-one consoles support out-of-band dead-node management functions, BIOS- and OS-level console redirection, and node health monitoring and management.

Selecting appropriate cluster management tools

The Dell OpenManage suite enables administrators to monitor and manage HPC clusters remotely, centralizing computing resources and streamlining management tasks to support enhanced accessibility. This software is also designed to help reduce staging, operational maintenance, and downtime and to enhance computing yield—with the ultimate goal of lowering total cost of ownership. The Dell OpenManage Installation and Server Management bundle consists primarily of Dell OpenManage Server Administrator (OMSA), which can be accessed through the CLI, Web browser–based graphical user interface (GUI), and Simple Network Management Protocol (SNMP) interface. When the GUI is not used, the corresponding Web server service can be shut down to save memory space for applications. The Dell OpenManage Service and Diagnostic Utilities bundle includes systems software extraction utilities, device drivers, and other utilities, and the Dell PowerEdge Updates bundle (designed for change management) includes the Software Update Utility and the Dell Update Package. The Dell OpenManage Deployment Toolkit (DTK) is a set of utilities that can configure nodes in a pre-OS state and provides scripted installation to configure a large number of compute nodes.

Apart from the Dell OpenManage suite, open source tools and utilities such as Ganglia, OpenIPMI, and IPMItool can be used effectively in conjunction with the Dell management framework to satisfy management requirements for some HPC environments. Platform OCS is a comprehensive cluster computing software stack designed to simplify software deployment and ease management of large-scale Linux-based clusters; this software stack and the individual tools packaged with it, such as Ganglia and Clumon, are discussed in more detail in the "Configuring and managing large-scale HPC clusters" section in this article. OpenIPMI is an IPMI software-level implementation to allow full access to IPMI on a server. IPMItool is a utility for managing and configuring devices that support IPMI 1.5 or 2.0, and also includes a CLI for BMC management.

Figure 1 shows a typical Dell HPC cluster architecture with public (external) access through NIC2, the out-of-band management channel through the Dell Remote Access Controller 5 (DRAC 5)

[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.

fabric, and the in-band management through the cluster and administration fabric over NIC1. NIC1 also supports out-of-band management through the BMC.

Dell OpenManage IT Assistant

Dell OpenManage IT Assistant is a centralized management console that provides administrators with a consolidated view of cluster status. Additionally, IT Assistant receives alerts from both OMSA and the BMC on a managed compute node when a server component is in a warning or critical state. Administrators can create alert filters in IT Assistant to isolate low-priority alerts that do not require an administrator's attention. To react to certain alert notifications automatically, administrators can set up alert actions on IT Assistant to execute commands to repair or mitigate the problem when a corresponding alert is received, or e-mail or page the administrator for action to help prevent parallel job failure from a hardware problem in advance.

Administrators can use IT Assistant as a central control point, and configure it to execute commands, perform device control, and update software. These features can help administrators efficiently manage and update cluster nodes remotely and in parallel—for example, by performing cluster firmware updates simultaneously based on a central repository on IT Assistant. Administrators can also export hardware status information from the IT Assistant database to other management consoles or tools to meet specific requirements.

BMC Management Utility

BMCs on ninth-generation Dell servers are IPMI 2.0 compliant, which enables administrators to monitor, configure, and manage systems remotely through an out-of-band fabric. The BMC monitors system status through sensors on the system board. When a sensor value exceeds a preset critical threshold, the BMC logs the alarm to the IPMI system event log for future reference. Meanwhile, administrators can configure alerts (Platform Event Traps) populated with event data and have them sent to a specified destination IP address.

Dell provides several utilities and programs for accessing and managing servers through the BMC, including the BMU, the Remote Access Configuration Utility, the DTK, OMSA, and CLI tools. To further enhance systems management security, the BMC provides an option for setting up encryption keys to encrypt IPMI sessions. IT Assistant can also access discovered BMCs through an out-of-band fabric.

Two BMC functions are frequently used to monitor and control cluster nodes: IPMI Shell (ipmish) and SOL.

Ipmish. The ipmish command-line utility is designed to access and manage remote systems through LANs or serial ports. To help improve the cost-effectiveness and management efficiency of serial fabric, administrators can enable ipmish over LAN to establish connections to remote managed systems using the existing Ethernet cluster administration fabric. Before using this tool, administrators must configure the BMC using the BMC setup menu or the Platform OCS BMC/BIOS configuration module developed for Dell PowerEdge servers. Ipmish allows administrators to check the current status of platform sensors, read system event logs, power manage servers, and so on. For example, to power down a hung node, administrators can issue the following command from a management station:

```
# ipmish -ip 80.255.255.251 -u root -p cal
    power off
```

SOL. The SOL Proxy utility allows administrators to view and change a cluster node's BIOS and BMC settings without an OS on the node, enable console redirection, and reboot a node from a management station. SOL Proxy is frequently used in HPC cluster environments as a remote console and for dead-node management. To use BIOS-level console redirection features, administrators must modify the serial port settings in the cluster node's BIOS in addition to configuring the BMC for ipmish: the Serial Communication option should be set to "On with Console Redirection via COM2," and the Failsafe Baud Rate value should be 115200. Administrators can modify these BIOS settings using the DTK or OMSA. If administrators need to set up OS-level console redirection, they must modify particular OS configuration files; for details, see the BMU user's guide or the Platform OCS BMC/BIOS configuration module developed for Dell PowerEdge servers.

To initiate multiple SOL sessions to access different cluster nodes simultaneously, administrators can configure the SOL Proxy server to support up to 20 SOL sessions. This setting can be changed in the solproxy.cfg file by specifying `max_session` from 1 to 20. To connect and use SOL Proxy, administrators must run a Telnet command with the SOL Proxy server IP address and port number in a command-prompt window—for example, `telnet localhost 623`. At the prompt, administrators can then enter the username, password, and IPMI encryption key. Once authentication is complete, administrators can activate console redirection by selecting the Activate Console Redirection option in the main menu. After console redirection begins, the text console of the remote managed system is redirected to the management station. Administrators can use the escape character sequence `~.` to terminate console redirection and return to the top-level menu.

Dell OpenManage Server Administrator

OMSA is a systems management application developed for Dell PowerEdge servers that includes both a GUI and a CLI. On Dell HPC cluster nodes, OMSA is usually installed across the cluster by using Platform OCS or other methods documented in the OMSA user's guide.


AMD 64
Opteron
The AMD Opteron™ processor provides industry-leading performance.
AT LAST, A LESS COMPLEX WAY TO BUILD YOUR EMPIRE.
INTRODUCING THE NEW DELL™ POWEREDGE™ 6950 SERVER.
Meet the new 4-socket Dell PowerEdge 6950 server. It's ideal for your big business, since it's primed to handle all your mission-critical database and virtualization apps. And Dell tests leading operating systems and applications on its servers to help ensure seamless productivity. So the sky is no longer the limit. Business solutions designed with one company in mind. Yours.
DELL
Purely You
See the Dell difference at
www.dell.com/poweredge
800-BUY-DELL
Dell cannot be responsible for errors in typography or photography. Dell, the Dell logo and PowerEdge are trademarks of Dell Inc. AMD, the AMD Arrow logo, AMD Opteron and combinations thereof, are trademarks of Advanced Micro Devices, Inc. ©2007 Dell Inc. All rights reserved.

To start an OMSA GUI session on a remote system, administrators can open a Web browser and type the host name or IP address of the managed node followed by the assigned port number (1311 by default)—for example, "https://*hostname*:1311" or "https://*IP address*:1311." Figure 2 shows the OMSA GUI after username and password authentication. The tabs at the top of the screen display available action categories for the selected object based on the administrator's access privileges. The data area displays detailed information for the selected system tree object and action tab.

The OMSA CLI enables administrators to use a command prompt to perform most of the systems management tasks available through the GUI—such as checking system status, changing BIOS and BMC settings, and configuring component probe warning threshold values. This CLI is also frequently used in HPC cluster environments to toggle firmware settings such as Intel Hyper-Threading Technology to help evaluate performance. Administrators can use the `omconfig` command to modify systems management settings, and the `omreport` command to query system component properties and status. For example, to enable SOL on a BMC, administrators can enter the following command:

```
omconfig chassis remoteaccess config=serialoverlan
   enable=true
```

To view the current SOL configuration on a BMC, administrators can use the following command:

```
omreport chassis remoteaccess config=serialoverlan
```

In Dell HPC clusters, the OMSA CLI provides flexibility for administrators to write scripts to customize OMSA functions for remote cluster nodes from a single management station.

DRAC 5–based out-of-band management

The DRAC 5 is designed to provide remote management capabilities, crashed system recovery, and power control functions for Dell PowerEdge servers. This controller has its own microprocessor and memory and is powered by the host system in which it is installed. For security, the DRAC 5 subsystem provides role-based authority, enabling each administrator to configure specific privileges, user IDs, and passwords through the Web browser–based GUI, which supports 128-bit Secure Sockets Layer (SSL) encryption, or the racadm CLI. Additional security features of the DRAC 5 include session time-outs, configurable IP ports, Secure Shell (SSH) logins, login failure limits per IP address, and a limited IP address range for clients connecting to it. The DRAC 5 should be used when graphics-mode console redirection, virtual media, role-based security, or a dedicated out-of-band fabric are required.

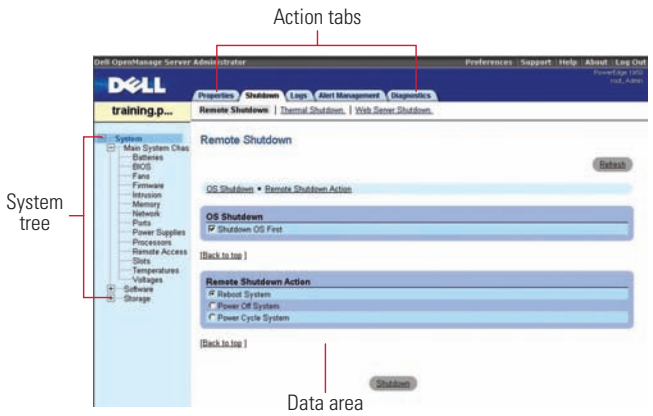


Figure 2. Dell OpenManage Server Administrator GUI

Configuring and managing large-scale HPC clusters

Platform OCS, a comprehensive cluster computing software stack based on software developed by the San Diego Supercomputer Center, is designed to simplify the deployment and management of large-scale Linux-based clusters. It can help administrators carry out deployment, maintenance, and management tasks on an HPC cluster. Platform OCS management tools include Ganglia and Clumon for monitoring, Platform Lava for job scheduling and management, Anaconda and kickstart for hosting images, Cluster Top for sensing the precise activity of cluster nodes, and the 411 user management system.

Ganglia is a scalable, distributed open source cluster monitoring system based on a hierarchical design that uses widespread technologies such as XML for data representation. It provides configurable cluster performance monitoring and can be configured to monitor cluster hardware health. Clumon is an open source cluster job state and resource utilization monitoring tool developed for Linux-based clusters at the National Center for Supercomputing Applications, and provides an overview of the current state of the cluster. It also provides a common structure and single point of access so that the many users and administrators do not overload the cluster with redundant requests for the same information. Platform Lava is a job scheduler designed for cluster resource management and job management.

Platform OCS is Roll based in that administrators can create Rolls to contain domain-specific software and applications; third-party vendors can customize Rolls to meet specific requirements. The Dell Roll contains automated scripts that administrators can use to configure the BMC and BIOS on supported Dell PowerEdge servers using the OMSA agent and its CLI utility. Using this Roll, cluster administrators can automate BMC and BIOS configuration for compute nodes in an HPC cluster, which can help significantly reduce deployment time. Administrators can use the OMSA omconfig utility to configure other components of the

BIOS, including console redirection. For PowerEdge SC platforms, administrators can use open source utilities such as IPMItool and OpenIPMI to automate BMC configuration during installation. Platform OCS also includes CLI utilities like cluster-ps, cluster-fork, and cluster-kill for performing cluster-wide operations (or operations on selected systems) from the command line.

Tools for each phase of the HPC cluster life cycle

The life cycle of a large-scale HPC cluster can be divided into several phases, with different management tasks at each stage:

- **Deployment or pre-OS phase:** Requires remote power management (through the BMC, Dell OpenManage IT Assistant, or DRAC) and remote BMC and BIOS setup (using Platform OCS or the DTK); utility partitions on hard drives are usually activated to diagnose and stage large-scale cluster hardware
- **Operational phase:** Involves cluster health monitoring through both in-band and out-of-band mechanisms, troubleshooting, self-healing (automated fan control), and auto-recovery (dead-node automated reboot), which are provided by IT Assistant, the DRAC, the BMU, and OMSA; also includes job scheduling, job monitoring, and resource management through Ganglia, Clumon, and Platform Lava
- **Change-management or maintenance phase:** Requires IT Assistant, the Software Update Utility, and the Dell Update Package

Dell OpenManage software components enforce security through role-based authority, allowing specific privileges for each user, 128-bit SSL encryption, and session time-out configuration using the GUI or CLI. Java Secure Socket Extension is used by supported operating systems to protect user credentials and other sensitive data.

Customizing hardware management for HPC clusters

Tools such as Dell OpenManage Server Administrator, Dell OpenManage IT Assistant, the Dell OpenManage Deployment Toolkit, the BMC Management Utility, the DRAC 5, and the Server Update Utility, along with Platform OCS, Linux services, Ganglia, Clumon, Platform Lava, and other software, can help provide an effective suite of utilities to configure, monitor, and manage HPC cluster environments. After evaluating the capabilities of each subsystem and its management requirements, administrators can utilize the strengths of each component to design a comprehensive management strategy for their specific HPC cluster environment.

Arun Rajan is a systems engineer in the Scalable Systems Group at Dell. His current interests and responsibilities include HPC cluster management, cluster computing packages, performance benchmarking, and product development. He has a B.E. in Electronics and Communications Engineering from the National Institute of Technology, Tiruchirappalli, and an M.S. in Computer and Information Science from the Ohio State University.

Tong Liu is a systems engineer in the Scalable Systems Group at Dell. His current research interests are HPC cluster management, high-availability HPC clusters, and parallel file systems. Tong serves as a program committee member of several conferences and working groups on cluster computing. Before joining Dell, he was an architect and lead developer of High Availability Open Source Cluster Application Resources (HA-OSCAR). Tong has an M.S. in Computer Science from Louisiana Tech University.

Yung-Chin Fang is a senior consultant in the Scalable Systems Group at Dell. He specializes in HPC systems, advanced HPC architecture, and cyberinfrastructure management. Yung-Chin has published dozens of conference papers and articles on these topics. He also participates in industry-standards organizations, academic conferences, and HPC cluster–related open source communities as a Dell representative.

Garima Kochhar is a systems engineer in the Scalable Systems Group at Dell. She has a B.S. in Computer Science and Physics from Birla Institute of Technology and Science, Pilani, and an M.S. in Computer Science from the Ohio State University.

Ron Pepper is a systems engineer and adviser in the Scalable Systems Group at Dell. He works on the Dell HPC Cluster team developing grid environments. Ron attended the University of Wisconsin at Madison, where he worked on a degree in Computer Science; he is continuing his degree at St. Edward's University.

FOR MORE INFORMATION

Dell systems management documentation:
support.dell.com/support/systemsinfo/documentation.aspx?~cat=6&~subcat=111

Ali, Rizwan, Rinku Gupta, Garima Kochhar, and Bill Bryce. "Platform Rocks: A Cluster Software Package for Dell HPC Platforms." *Dell Power Solutions,* November 2005. www.dell.com/downloads/global/power/ps4q05-20050227-Ali.pdf

Kochhar, Garima, Rizwan Ali, and Arun Rajan. "Configuring the BMC and BIOS on Dell Platforms in HPC Cluster Environments." *Dell Power Solutions,* November 2005. www.dell.com/downloads/global/power/ps4q05-20050222-Kochhar.pdf

Rajan, Arun, Tong Liu, Yung-Chin Fang, and Saeed Iqbal, Ph.D. "Secure HPC Cluster Management with Ninth-Generation Dell PowerEdge Servers." *Dell Power Solutions,* November 2006. www.dell.com/downloads/global/power/ps4q06-20060425-Fang.pdf

Using the Dell OpenManage Deployment Toolkit to

Automate Configuration of HPC Cluster Nodes

High-performance computing clusters comprising hundreds of nodes require one-to-many management utilities to configure the nodes with customized BIOS, baseboard management controller, and Dell™ Remote Access Controller settings and to deploy Dell Utility Partitions in a reasonable amount of time. Administrators can use the Dell OpenManage™ Deployment Toolkit to perform these one-to-many operations across multiple nodes simultaneously.

BY SHRANKHLA UPADHYAY, MEGHANA BHAT, BARIS GULER, AND ARUN RAJAN

Related Categories:

High-performance computing (HPC)

System deployment

Systems management

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

High-performance computing (HPC) clusters combine several computer systems with a high-speed interconnect to achieve raw output comparable to traditional supercomputers but at a lower cost. HPC clusters divide large individual tasks among several computers, regulate the processes for high throughput, and present the results to a client as if processed by a single system. Such clusters can handle a broad range of numeric- and data-intensive applications efficiently and are finding favor with organizations in many fields because of their cost-to-performance ratio.

HPC cluster architecture features a parallel computing system consisting of one or more master nodes and one or more compute nodes interconnected by a private network system. The master node serves to synchronize the compute nodes and maintain general cluster health; the compute nodes divide and process the computational tasks. The size of a cluster can range from a few hundred to a few thousand nodes, depending on the organization's performance requirements.

The Dell OpenManage suite offers command-line interface utilities to configure BIOS, baseboard management controller (BMC), Dell Remote Access Controller (DRAC), and RAID settings across supported cluster nodes through the network using the Preboot Execution Environment (PXE) or a CD. However, these utilities can be used only after an OS is deployed, and preconfiguring system settings before OS deployment can often be beneficial. Traditional methods for pre-OS configuration include using KVM (keyboard, video, mouse), console redirection, or a crash cart to manually modify the system settings through the boot menu options one server at a time—a process that can be time-consuming, repetitive, and tedious.

Administrators can use the Dell OpenManage Deployment Toolkit (DTK) to perform one-to-many pre-OS operations across multiple Dell PowerEdge™ servers simultaneously. The DTK consists of command-line utilities to configure system features and provides sample scripts to perform common deployment tasks. The DTK utilities can be integrated with the master node's cluster

management software and then pushed onto individual client nodes without configuring each node individually. This article describes how the DTK can be used in a Linux® OS environment for typical configuration tasks performed on HPC clusters before the deployment of the cluster package as well as during maintenance tasks. It also discusses techniques for configuring BIOS, BMC, and DRAC settings and installing the Dell Utility Partition; these techniques can be scaled to include other DTK configuration features.

Using the Dell OpenManage Deployment Toolkit in HPC clusters

The life cycle of a typical HPC cluster can be divided into several phases, with different management tasks in each phase. The deployment phase involves remote power management, which is provided by the BMC and requires setting up parameters in the BMC options menu such as IP address, netmask, user ID, and password. The operational phase involves cluster health monitoring, which requires both in-band and out-of-band mechanisms be in place. To enable these mechanisms, BMC options such as username, password, and LAN alert destination and BIOS options such as boot sequence, snoop filter, Demand-Based Switching, logical processor, and console redirection settings must be configured.

On ninth-generation servers, the DTK allows administrators to modify most configuration parameters of the remote system, specifically the BIOS, BMC, and DRAC 5 settings. Administrators can use the DTK syscfg tool to capture and replicate these settings. Alternatively, administrators can use an automated script to capture a complete list of BIOS and BMC settings and use this list as a comparison when troubleshooting problems such as failed nodes and performance declines. The DTK can help provide a safe preinstallation OS environment for performing necessary operations during the deployment of an HPC cluster, helping reduce the time spent manually configuring each node in the cluster. Figure 1 outlines some DTK features and associated commands relevant to cluster deployment.

Setting up the Dell OpenManage Deployment Toolkit environment

The DTK is available in two versions: a Microsoft® Windows® Preinstallation Environment (WinPE) self-extractable zip file and a Red Hat® Enterprise Linux 3 ISO image. Administrators can extract the WinPE tools and scripts obtained from the zip file and build them into a WinPE image. Administrators can also extract the tools and scripts from the Linux ISO image and use them in a customized ISO image, or use the image provided with the DTK. The information provided in this article refers to the Linux version of the DTK.

Administrators can set up the DTK environment in a cluster using the PXE mechanism available on the master node. The PXE environment allows servers (in this case, the compute nodes) to boot over a network connection. As each compute node boots into the lightweight Linux environment provided by the DTK through PXE from the master node, it executes the automated scripts that perform the required operations on the system.

Installation prerequisites are as follows:

- DTK 2.*x* CD (created by downloading the DTK ISO image from support.dell.com and burning the image on a CD)
- Master node running Linux or a Linux-based OS
- Compute nodes connected to the master node and ready to install

Setting up the DTK environment on the master node

A wide variety of cluster management software is available, and administrators can use the DTK to configure basic BIOS, BMC, DRAC, and RAID settings on different types of Dell Linux clusters. Administrators must configure the Trivial FTP (TFTP) server on the master node; the directory used to serve files is typically the /tftpboot directory. The DTK kernel file (SA.1) and ramdisk file (SA.2), available with the DTK download, must be copied to the master node /tftpboot directory.

After the DTK CD is mounted on the server, the files can be copied to the /tftpboot directory using the following commands:

```
mount /dev/cdrom /mnt/cdrom
cp /mnt/cdrom/isolinux/SA.1 /tftpboot/SA.1
cp /mnt/cdrom/isolinux/SA.2 /tftpboot/SA.2
umount /mnt/cdrom
```

Administrators must also create a folder on the Network File System (NFS) share of the master node. In the example scripts shown in this article, /export/home/install/dtk carries the default sample scripts provided with the DTK as well as any custom scripts written for specific server configurations.

Finally, administrators must place an entry for the DTK in the standard master node PXE configuration file for all the compute nodes to obtain the DTK kernel and ramdisk images over PXE. Figure 2 shows an example PXE configuration file. The IP address of the master node (for the private network of the cluster) is set to 10.1.1.1. Additional examples for different types of shares are available in the isolinux.cfg file provided with the DTK Linux ISO image.

Installing the Dell Utility Partition

The Dell Utility Partition is a bootable partition on the hard drive that contains diagnostic and system utilities. The partition is typically between 32 MB and 60 MB. Installing the partition requires the latest Dell Utility Partition images, which are provided by Dell OpenManage Server Assistant (available on the Dell Installation and Server Management CD that ships with PowerEdge servers).

DTK option	Get example	Set example	Description
Asset	`syscfg --asset`	`syscfg --asset=node1`	Sets the asset tag value, which can only be assigned programmatically; Dell OpenManage Server Assistant and the DTK enable administrators to assign this information
Boot sequence	`syscfg --bootseq`	`syscfg --bootseq=3,2,1`	Sets the node boot sequence

LCD	`syscfg --lcd`
	`syscfg --lcd1`
	`syscfg --lcd2`
		`syscfg --lcd=user`
		`syscfg --lcd1="Change HDD"`
			Sets the PowerEdge server LCD to reflect information about the server

Intel® Hyper-Threading Technology	`syscfg --logicproc`	`syscfg --logicproc=enable`	Enables or disables the logical processor
Power button	`syscfg --powerbutton`	`syscfg --powerbutton=disable`	Enables or disables the node power button
USB ports	`syscfg --usbports`	`syscfg --usbports=disable`	Enables or disables USB ports
Adjacent cache line prefetcher	`syscfg --adjcacheprefetch`	`syscfg --adjcacheprefetch=enable`	Enables or disables cache line prefetch settings to help enhance system performance
Hardware prefetcher	`syscfg --hwprefetcher`	`syscfg --hwprefetcher=disable`	Enables or disables hardware prefetch settings to help enhance system performance

LAN	`syscfg lancfgparams`	`syscfg lcp`
		`--alertdest=1`
		`--destipaddr=192.168.0.101`
		`--nicselection=dedicated`
		`--dnsserver1=192.168.0.100`
		`--dnsracname=rac-dell`
			Configures and displays LAN configuration parameters

LAN users	`syscfg lanuseraccess`	`syscfg lua`
		`--userid=2`
		`--usrprivlmt=administrator`
			Configures the privilege level and channel accessibility associated with a specified user ID for the LAN channel

Platform Event Filter (PEF)	`syscfg pefcfgparams`	`syscfg pcp`
		`--filter=powerfail`
		`--filteralert=enable`
			Configures and displays PEF configuration parameters

Demand-Based Switching (DBS)	`syscfg --dbs`	`syscfg --dbs=enable`	Enables or disables DBS; can help minimize power and cooling requirements
Network interface card (NIC)	`syscfg --embnic1`	`syscfg --embnic1=on` `syscfg --embnic2=onnopxe`	Configures the embedded NIC ports to enable or disable PXE and turn NIC ports on or off
Snoop filter	`syscfg --snoopfilter`	`syscfg --snoopfilter=enable`	Enables or disables the snoop filter option on supported systems

Figure 1. System parameters configurable with the DTK

```
label dtk
kernel SA.1
append initrd=SA.2 ramdisk_size=36817 Stage3_type=cdrom DEBUG=0 quiet
ide=force_255_heads share_type=nfs share_location=10.1.1.1:/export/home/install/dtk share_script=dtk.sh
   share_opts=nolock
```

Figure 2. Example PXE configuration file


```
svctag='syscfg --svctag | awk -F= '{print $2}''
sysname='syscfg --sysname | awk -F= '{print $2}' | cut -c11-14'
./upinit.sh --overwrite -disk=/dev/sda -size=32 -file=../../systems/UP/2800/upimg.bin
```

Figure 3. Sample dtk.sh script

```
syscfg --bootseq=cdrom.emb.0,floppy.emb.0,hdd.emb.0,nic.emb.1
syscfg --logicproc=disable
syscfg --snoopfilter=enable
syscfg --dbs=enable
syscfg --embnic1=on
syscfg --embnic2=onnopxe
```

Figure 4. Sample dtk.sh script snippet for BIOS configuration

```
syscfg lancfgparams --ipaddress 10.0.0.3 --subnetmask 255.0.0.0
syscfg lancfgparams --alertdest=1 --destipaddr=10.0.0.20
syscfg lancfgparams --nicselection=dedicated
```

Figure 5. Sample dtk.sh script snippet for BMC and DRAC configuration

The compute nodes must install the partition before installing the HPC cluster management software. Therefore, during installation, the preferred label in the PXE configuration file, `default`, under the master node /tftpboot/pxelinux/pxelinux.cfg folder is changed to `dtk`, causing the compute nodes to boot from the DTK image for the installation of the Dell Utility Partition. Administrators can make the compute nodes boot from the master node PXE setup by pressing F12 as the nodes start to boot.

The dtk.sh script, a sample of which is shown in Figure 3, is used to install the Dell Utility Partition; this script should be placed in the NFS share at /export/home/install/dtk. Once the Dell Utility Partition installation is complete on all compute nodes, the master node label is changed back to the original label, and the standard cluster installation can proceed normally.

Capturing and modifying system configuration settings

Administrators can capture or modify BIOS, BMC, and DRAC settings using DTK commands such as those shown in Figure 1. To capture all system configuration settings accessible by the DTK, administrators can use the syscap.sh script, available at /export/home/install/dtk/template/scripts, as a default template. Figures 4 and 5 show example script snippets that administrators can incorporate into the main dtk.sh automated script file.

Configuring cluster nodes with the Dell OpenManage Deployment Toolkit

The Dell OpenManage Deployment Toolkit can help administrators perform a wide range of server configuration tasks on large HPC clusters. Using DTK features and its preconfigured scripts in a pre-OS environment can help simplify the setup and configuration of critical server parameters in both homogenous and heterogeneous HPC clusters, with the ultimate goal of helping improve data center deployment, management, and administration, and eventually total cost of operation.

Shrankhla Upadhyay is an engineering analyst in the Scalable Systems Group at Dell. Her current interests include HPC cluster management, operating systems, and cluster file systems. She has a B.E. in Computer Science from Birla Institute of Technology and Science, Pilani.

Meghana Bhat is a senior engineer in the Deployment team in the Dell OpenManage Product Group. Her areas of interest include open source software and standards. She has worked on Dell OpenManage Server Assistant and is currently working on the Dell OpenManage Deployment Toolkit. Meghana has a B.E. in Computer Science from Vishwakarma Institute of Technology, Pune.

Baris Guler is an HPC application specialist in the Scalable Systems Group at Dell. His current research interests are parallel processing, diskless HPC clusters, performance benchmarking, reservoir engineering and simulation, and numerical methods. Baris has a B.S. in Petroleum and Natural Gas Engineering (PNGE) from the Middle East Technical University and an M.S. in PNGE from the Pennsylvania State University.

Arun Rajan is a systems engineer in the Scalable Systems Group at Dell. His current interests and responsibilities include HPC cluster management, cluster computing packages, performance benchmarking, and product development. He has a B.E. in Electronics and Communications Engineering from the National Institute of Technology, Tiruchirappalli, and an M.S. in Computer and Information Science from the Ohio State University.

Creating Scalable Compute Clusters

Using the Sanbolic Melio File System and LaScala Volume Manager

High-performance compute clusters can provide the performance and scalability necessary for processing-intensive applications. This article discusses using the Sanbolic® Melio® File System and Sanbolic LaScala® volume manager to create cost-effective, scalable clusters based on Dell™ servers and storage and the Microsoft® Windows® OS.

BY WILLIAM STEVENSON

Related Categories:

File systems

High-performance computing (HPC)

Microsoft Windows

Sanbolic

Storage

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

For enterprises using processing-intensive applications such as video processing or mathematical modeling programs, cost-effectiveness and scalability can be essential. Building a shared data processing cluster with standards-based Dell PowerEdge™ servers and Dell PowerVault™ and Dell/EMC storage running Microsoft Windows OS–based applications and a symmetrical clustered file system can provide a cost-effective, scalable storage and compute platform by enabling administrators to add compute nodes or storage arrays as needed.

Compute cluster architecture

Block-level storage and a clustered file system are key elements of a compute cluster architecture. Using block-level storage enables the compute cluster to deliver the necessary bandwidth even in large clusters processing large image or video files, and a clustered file system enables all cluster servers to have shared read/write access to the storage. A clustered file system can also provide concurrent file access for a large number of clients by using a network attached storage (NAS) gateway to share access to the same storage area network (SAN) volumes.

This architecture is readily adaptable to video processing or mathematical modeling applications.[1] Figure 1 shows a typical example for a video rendering application, in this case Autodesk Maya, along with the Rush job scheduler to distribute rendering jobs across the cluster. This type of configuration is typically appropriate for midsize compute clusters of 4 to 64 nodes.

The clustered file system and volume manager for this configuration are provided by the Sanbolic Melio File System (Melio FS) and Sanbolic LaScala clustered volume manager, which are installed on both NAS heads and on each cluster node. Melio FS is used to format the volumes and provide shared file-level access to data on the SAN; LaScala enables volumes to be striped across multiple storage arrays to aggregate storage bandwidth. The Microsoft Windows Storage Server 2003 OS–based NAS heads and the Microsoft Windows Server® 2003 OS–based cluster nodes all have concurrent read/write access to Melio FS over the Fibre Channel SAN.

The Microsoft Windows 2000 and Windows XP OS–based workstations access the NAS heads using the Common Information File System (CIFS) over the LAN,

[1] Although currently only a limited number of applications are designed to use a shared data compute infrastructure, this number is expected to increase as Microsoft develops its partner community around Microsoft Windows Compute Cluster Server 2003.

and the NAS heads write data to Melio FS on the storage arrays. This data is then available through the Fibre Channel SAN to the compute nodes, which process the data using Windows-based applications and write results back to the SAN. The workstations can then access the results from the NAS heads using CIFS.

This type of configuration is designed to be scalable: the LaScala volume manager allows the dynamic expansion of storage volumes as administrators add storage arrays. If administrators need to increase the computational power, they can also add servers to the cluster, which can then have immediate access to the shared data written to Melio FS. The configuration can use both Internet SCSI (iSCSI) and Fibre Channel storage networking in a single installation with shared access to SAN volumes, but administrators should match the performance of the storage network infrastructure to the application requirements on each server.

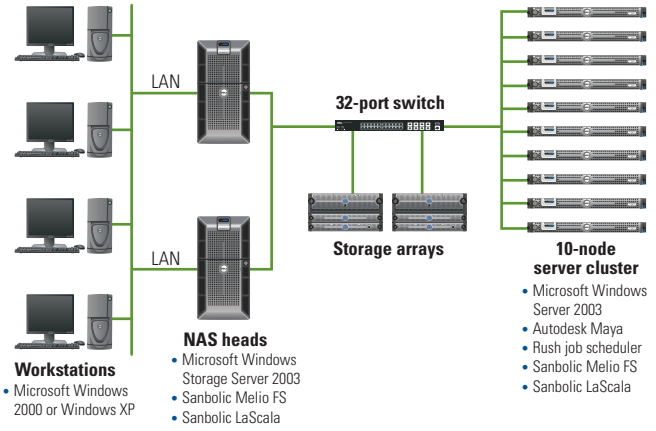


Figure 1. Typical compute cluster configuration for a video rendering application using block-level storage, Sanbolic Melio File System, and Sanbolic LaScala

Clustered file system

Melio FS and LaScala are designed for shared SAN environments. LaScala allows administrators to centrally manage the entire SAN storage pool, with access to volumes assigned using native Windows access control list (ACL) security. The file system supports the Microsoft Active Directory® directory service and enforces user account access rights across the storage cluster. Melio FS appears to each workstation or server as a local volume, while maintaining cache coherency among all cluster nodes and managing distributed locking to permit concurrent read/write access to files from multiple servers or workstations.

This architecture is straightforward to install, and requires no proprietary hardware or complex metadata server configuration. Each server connected to the storage cluster requires a Fibre Channel or iSCSI connection to the storage array and a TCP/IP connection for lock traffic. Administrators can typically install the file system and volume manager on each SAN-connected Windows-based server or workstation in a few minutes. Any Microsoft application that supports distributed processing can be used with this architecture.

Cluster performance and scalability

A typical rendering node requires approximately 25 MB/sec of storage bandwidth, depending on the type of rendering being performed. Therefore, a 10-node rendering cluster has an aggregate storage bandwidth requirement of about 250 MB/sec, and a cluster larger than that has proportionately higher bandwidth requirements. Shared Fibre Channel storage can readily meet these requirements. For large clusters, administrators can aggregate additional storage arrays into a single shared volume using LaScala, which efficiently combines the storage bandwidth of multiple arrays while maintaining the native block size of the application.

The example configuration shown in Figure 1 uses compressed standard-definition video, and can adequately support the workstations' storage bandwidth requirements using a storage connection through a NAS head. For high-performance applications such as those for high-definition video editing (which can require 165 MB/sec for uncompressed 1080i content), workstations can be connected directly to the shared SAN volume using Fibre Channel. Melio FS appears as a local drive, and the workstations have concurrent read/write access to the data on the volume at Fibre Channel speed.

Melio FS introduces a small amount of overhead, which varies with the block size of the application. Performance for large image or video files is typically equal to or better than using a local NT File System (NTFS) format, because lock overhead is minimal relative to other performance optimizations. Small-block applications may experience 10–15 percent performance overhead in storage bandwidth for each node.

Cost-effective platform for high-performance clusters

Although Windows applications currently supporting clustered processing hardware tend to be limited to seismic imaging and video rendering programs, such support should soon be extended to include a broader group of computer-aided design and engineering, life sciences, and financial modeling programs. Compute clusters with clustered file systems and block-level storage connections, when built with standards-based Dell servers and storage, can provide a cost-effective platform for these types of processing-intensive applications, and the increasing availability of multi-core technology and iSCSI storage connections over 10 Gigabit Ethernet infrastructures should serve to enhance both the performance and scalability of these clusters in the future.

William Stevenson is the executive chairman of Sanbolic and has been with the company since 2001. He has an A.B. in Economics from Princeton University and studied at the University of Cologne with a Fulbright Scholarship.

FOR MORE INFORMATION

Sanbolic:
www.sanbolic.com

Linux Device Naming

Like other operating systems, the Linux® OS must discover, enumerate, and assign default names to hardware devices—tape drives, network adapters, USB devices, and the like—before these devices can be used. This article discusses methods for Linux device enumeration, how enumeration relates to default assigned names, and how administrators can modify these names.

BY MATT DOMSCH AND AHMAD ALI

Related Categories:

Linux

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

Names provide context for objects and create explicit or implied relationships between objects. For computers, however, an object name may have little to do with its underlying characteristics.

Consider a real-life anecdote of a child born at a hospital in Austin, Texas. The hospital gave the baby the name "BGC Dom." This name appeared on her wristbands and bassinet cart, the nurse's duty station whiteboard, and hospital paperwork. The hospital had created a naming convention designed to be easily recognizable to the staff, unambiguous (given an average of fewer than 20 babies at the hospital at a time), and protective of the mother's and baby's privacy: "BG" stood for "baby girl," "C" was the first letter of the mother's first name, and "Dom" was the first three letters of the mother's last name. (The convention also included extensions for identical twins.) Because her parents did not want to call her "BGC Dom" throughout her life, they filled out other paperwork to officially name her (involving the baby's globally unique footprints to identify her). However, even after the parents completed the official naming, the hospital continued to use the name they had assigned, "BGC Dom"—and yet no one was confused about who the baby was or who the mother was. This is the goal of any good naming system.

System device naming in the Linux OS is similar in several ways. Conventions are in place to assign default names to devices, but the default name may not be the only name a device has, and the device name can be changed to help ease identification and management of that device.

Understanding enumeration and naming

Enumeration is the process of moving through the system in some order: depth first, breadth first, by some list of subsystems, and so on. Each control entity (Linux kernel,

BIOS, and system administrator) performs its own device enumeration and naming. Each discovered device is assigned a default name—for example, in Linux, Ethernet devices become eth0, eth1, eth2, and so on, and SCSI and Serial ATA (SATA) disks become /dev/sda, /dev/sdb, and so on. The Linux default naming convention is simply the device type (eth or sd) plus an incrementing value based on the order in which devices are found. The BIOS naming is "interrupt 13h device 80h, device 81h," and so on, while the system administrator may name devices as "my boot disk," "my home directories," "my database," and so on.

Such default naming conventions can present several problems:

- The default device name assigned by the OS may bear no resemblance to the name assigned by the system BIOS or hardware manufacturer documentation, and it may have no relationship to eventual system usage.
- A given device name may change if the system topology or enumeration order changes.
- As one device name changes, another device may take the default name previously assigned to the first device (/dev/sdb may become /dev/sda, eth1 may become eth0, and so on).
- Administrators are accustomed to thinking that the default name is the correct name or the only name. If the name changes, the administrator will have an inaccurate conception of the system or device.
- Tools and configuration files that use default device names may behave incorrectly if the names change.

In an ideal world, each device would have some characteristic besides enumeration order to use for naming. Just as in the scenario described earlier in which the baby's unique footprints are registered for her birth certificate, many (but not all) devices have some form of unique identifier: for example, Ethernet devices have 48-bit globally unique Media Access Control (MAC) addresses, and disks have manufacturer serial numbers or other globally unique identifiers stored in their firmware. The naming system could have tables and algorithms for assigning consistent names to devices independent of their enumeration order.

In an ideal world, each device would have some characteristic besides enumeration order to use for naming.

Linux 2.6 kernels and the udev tool enable administrators to assign arbitrary names to devices following their own naming conventions. The Linux device enumeration and default naming process varies depending on the type of device, with conventions for Ethernet devices differing from those for disk devices.

Naming Ethernet devices

Default naming for Ethernet devices uses device enumeration order; enumeration order is based on the order in which network device driver modules are loaded, which in turn is based on the order in which those drivers are listed in /etc/modprobe.conf. If a system has one Ethernet device that uses the Intel® e1000 driver and another that uses the Broadcom tg3 driver, administrators can affect the enumeration order of these devices based on the order of entries in /etc/modprobe.conf.

Multiple Ethernet devices using the same driver can be more complicated than single devices for each driver. If the drivers in the preceding example are being used by multiple devices, the e1000 driver enumerates all the Ethernet ports it recognizes (for example, eth0, eth1, and so on), and then the tg3 driver, when loaded, enumerates those it recognizes (eth2, eth3, and so on). Discovery order is based on the system-wide PCI device list (which itself was enumerated in a particular order) and the list of PCI device IDs that the driver can recognize. This enumeration makes no distinction between LAN on Motherboards (LOMs) and add-in PCI cards, which can cause two LOMs to be enumerated backward compared to their BIOS names (NIC1 as eth1, NIC2 as eth0), and cause add-in cards to be assigned eth numbers that are lower than the LOM port names, leading to confusion for system administrators.

Because each Ethernet port has a unique MAC address, Linux can assign any name to a given MAC address. Administrators have traditionally used the `nameif` command and the related /etc/mactab file to assign logical names (public0, private1, subnet15, and so on) to each interface. Red Hat® Enterprise Linux 3 and 4, however, have configuration files in /etc/sysconfig/network-scripts/ifcfg-*devicename* that include a `HWADDR=`*xx*`:`*xx*`:`*xx*`:`*xx*`:`*xx*`:`*xx* value used to assign a device name to the Ethernet port based on the specified MAC address, overriding the default name assigned during device enumeration.

Novell® SUSE® Linux Enterprise Server (SLES) 9 and 10 both store configurations in files like /etc/sysconfig/network/ifcfg-eth-id-*xx*:*xx*:*xx*:*xx*:*xx*:*xx*, where the last part of the file name is the Ethernet port's MAC address. By default, the OS assigns eth0 to the first device found. In SLES 10, the udev tool maintains name persistence for Ethernet ports by renaming ports from the kernel default naming scheme. This tool establishes naming rules using files in /etc/udev/rules.d; it also has rules for adding Ethernet devices. The /lib/udev/rename_netiface script carries out this renaming. Administrators can choose to establish new rules or modify default ones. Each rule is a single line with keywords used to match or assign a value or perform an action. The "Using the udev tool for device naming" sidebar in this article provides some example rules.

Dell engineers have developed a script that assigns Ethernet device names: the LOMs are assigned names that match BIOS and documentation-defined order, and Ethernet ports on add-in cards are then named in increasing slot number order (the slot numbers are

USING THE UDEV TOOL FOR DEVICE NAMING

The udev tool provides device file management capabilities in Linux distributions based on the 2.6 kernel. It uses rules to dynamically maintain (create, remove, or rename) files for devices that are actually present in the system. It is based on sysfs and hotplug, and replaces devfs and static device node naming.

The default location for udev configuration files is /etc/udev. The main configuration file is /etc/udev/udev.conf. This file has information such as udev_root, udev_rules, permissions, and logging priority. The default udev_root location is /dev; the default udev_rules location is /etc/udev/rules.d.

The udev rules are stored in files under the udev_rules directory and are processed in lexical order. Every non-blank or non-comment line in a rule file is a rule. These rules contain one or more key value pairs separated by a comma. These are both match keys and assignment keys. The rules are applied when all match keys are matched; when no rules match, default device names are used.

Rules may name a device, add a symlink, or run a program to take other actions like renaming an Ethernet port. The key match operators are = and !=; the assignment operators are =, :=, and +=. Key value assignments can be changed by subsequent rules unless the final assignment operator, :=, is used.

Figures A–E show some example rules. *Note:* The exact syntax for udev rules has changed over time; for an OS distribution's specific udev syntax, see the distribution documentation.

```
SUBSYSTEM=="net", ACTION=="add", SYSFS{address}=="xx:xx:xx:xx:xx:xx",
    IMPORT="/lib/udev/rename_netiface %k eth0"
```

Figure A. Example udev rule that renames Ethernet ports at a given MAC address to eth0 with the help of an external script

```
KERNEL=="*[!0-9]", IMPORT{program}="/sbin/edd_id --export $tempnode"
KERNEL=="*[!0-9]", ENV{ID_EDD}=="?*",
    SYMLINK+="disk/by-id/edd-$env{ID_EDD}"
KERNEL=="*[0-9]", ENV{ID_EDD}=="?*",
    SYMLINK+="disk/by-id/edd-$env{ID_EDD}-part%n"
```

Figure B. Example udev rule that creates a symlink for disks and partitions using BIOS disk data, providing persistent names based on BIOS disk enumeration

```
BUS="scsi", PROGRAM="/sbin/udev.get_persistent_device_name.sh",
    NAME="%k" SYMLINK="%c{1+}"
```

Figure C. Example udev rule that adds a symlink with a helper program that provides path and UUID information

```
KERNEL=="sd*[!0-9]|sr*|st*", ENV{ID_SERIAL}=="",
    IMPORT{program}="/sbin/scsi_id -g -x -s %p -d $tempnode"
KERNEL=="sd*[!0-9]|sr*|dasd*[!0-9]", ENV{ID_SERIAL}=="?*",
    SYMLINK+="disk/by-id/$env{ID_BUS}-$env{ID_SERIAL}"
KERNEL=="sd*[0-9]|dasd*[0-9]", ENV{ID_SERIAL}=="?*",
    SYMLINK+="disk/by-id/$env{ID_BUS}-$env{ID_SERIAL}-part%n"
```

Figure D. Example udev rule using the SLES 10 syntax that adds symlinks for disks and all partitions

```
BUS=="scsi", ENV{ID_SERIAL}=="360045600007890012", NAME="BackupDisk4%n"
KERNEL=="sd*[!0-9]", ENV{ID_SERIAL}=="360045600007890012",
    SYMLINK+="location/Rack1Shelf2Disk4"
```

Figure E. Example udev rules that provide a persistent friendly name and symlink based on SCSI disk serial number

printed inside the system). This script is available at linux.dell.com/files/name_eths, and works on Red Hat Enterprise Linux and SLES 10.

Naming disk devices

Default naming for disk devices is also based on device enumeration order, which in turn is based on the order in which the device drivers are loaded. For standard IDE disks and CD/DVD drives, the OS discovers these drivers during kernel startup, before other kernel modules are loaded. IDE devices are named /dev/hda, /dev/hdb, and so on. Following the monolithic kernel startup, enumeration continues by loading the drivers listed in the initial ramdisk, which is generated from the order listed in /etc/modprobe.conf. SCSI, SATA, Serial Attached SCSI (SAS), Fibre Channel, and the various hardware RAID controller drivers are all modular, and thus are loaded from the initial ramdisk. Because these drivers are based on the SCSI midlayer, the OS names the devices /dev/sda, /dev/sdb, and so on.

Administrators should note that the Linux kernel discovery order has no relationship with the BIOS boot order—the disk that Linux calls /dev/hda or /dev/sda may not be the disk the BIOS uses to boot the OS. This is usually only a problem during OS installation, because administrators must ensure that the boot loader, /boot partition, and often the root (/) partition are on the disk the BIOS uses to boot the OS. Further complicating the problem is that Linux may not be able to ask the BIOS directly for its list of devices.

Dell engineers have developed a method to maintain device naming consistency for disks. By writing a system-wide unique signature into the master boot record (the first sector) of each disk in the system, the installation software can use this identifier to match the BIOS device list with the Linux device list. Early in the kernel startup process, each disk is read, and the signature value is stored in memory and later exposed through sysfs. Linux user-space tools such as the Red Hat Enterprise Linux 4 Anaconda installer and SLES 10 YaST (Yet Another Setup Tool) installer can then read the signatures from all the disks, compare them to the BIOS disk signatures shown in sysfs, and map the BIOS device names to the Linux device names.

After the system is fully installed, the standard method to mount the file systems in the proper place is to use file system labels in the /etc/fstab file. For example:

```
LABEL=/        /         ext3    defaults 1 2
LABEL=/boot    /boot     ext3    defaults 1 2
```

This example indicates that the ext3 file system with the / file system label should be mounted at the mount point /, and the /boot file system label should be mounted at the mount point /boot. As long as the file system labels in use are unique across the system, the file systems can then be mounted in the correct location. In Red Hat Enterprise Linux 3 and later, the boot loader configuration file can use this label to find the root file system. Administrators should pass `root=LABEL=/` on the kernel command line, although they must use an initial ramdisk for this to work, because the shell in the initial ramdisk resolves the requested label to the actual underlying disk partition and file system.

Using label-less file systems

Some file systems (such as, until recently, swap devices) and raw devices do not include useful file system labels, but administrators can use other methods to assign consistent names to these devices. For Red Hat Enterprise Linux 3, Dell has developed the devlabel tool, which uses unique SCSI serial numbers or universally unique identifiers (UUIDs), depending on the individual disk capability, and a configuration file to map the identifiers to the desired mount point. For Red Hat Enterprise Linux 4, SLES 9, and SLES 10, this capability has been superseded by udev. Administrators can set the udev rules to provide persistent symbolic links (symlinks) using UUID, path, serial number, or file system labels. For example udev rules, see the "Using the udev tool for device naming" sidebar in this article.

Divorcing enumeration from naming enables administrators to assign device names based on local logical usage and to change the physical topology of systems without affecting administrator expectations or system software.

Controlling Linux device naming

Although Linux assigns default names to all types of devices, these names may not be the most useful names. Fortunately, administrators can use a number of methods to change the names depending on their needs. Divorcing enumeration from naming enables administrators to assign device names based on local logical usage and to change the physical topology of systems without affecting administrator expectations or system software.

Matt Domsch is a Linux software architect at Dell and a member of the Fedora Project Board. Matt has a B.S. in Computer Science and Engineering from the Massachusetts Institute of Technology and an M.S. in Computer Science from Vanderbilt University.

Ahmad Ali is a member of the Dell Linux Engineering team. He has a B.S. in Electrical Engineering and an M.S. in Computer Engineering from Florida Atlantic University.

Upgrading BIOS and Firmware

on Linux-based Systems with Firmware-tools

Firmware-tools, an open source development project, aims to simplify the process of upgrading BIOS and firmware on Linux® OS–based systems. This article discusses the firmware-tools architecture, which delivers and installs these upgrades using the native Red Hat® Package Manager (RPM™) format and native Linux change-management frameworks such as Novell® ZENworks® Linux Management and the yum tool.

BY MATT DOMSCH AND MICHAEL E. BROWN

Related Categories:

Linux

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

BIOS and firmware updates can be implemented in various ways depending on the system software environment. These updates can be delivered on floppy disks that boot into a DOS-like environment; however, floppies are slow, limited in capacity, and usually require manual operation, making them difficult to deploy remotely or automatically. System vendors have also produced BIOS and firmware installers that work from within an OS.

Firmware-tools, a Dell-developed open source project, is designed to align the delivery and installation of BIOS and firmware updates with existing Linux packaging and distribution formats. This method may not yet be officially supported by all system vendors, but community participation in project development is welcome.

Understanding the firmware-tools architecture

The firmware-tools project builds on the capabilities of the existing Dell™ Update Packages for Linux—self-contained BIOS and firmware installers for Dell systems running Novell SUSE® Linux Enterprise Server or other Linux operating systems. Each package provides an inventory tool to discover current system information and compare it with the new payload version, an executable tool to install the payload on a flash memory chip, and the payload data. Dell also provides tools to integrate these packages into the Altiris®, Microsoft® Systems Management Server, and Novell ZENworks Linux Management Dell Edition change-management frameworks to help simplify deployment using these platforms.

The primary goal of the firmware-tools project is to integrate BIOS and firmware updates into the same change-management frameworks administrators already use to manage their operating systems, which can shorten the administrative learning curve and help improve efficiency. A secondary goal is to support additional operating systems (for example, OpenSUSE and Fedora) and system types (servers, workstations, desktops, and notebooks) as development resources permit. Finally, splitting the inventory, executable, and payload components into separate packages can help reduce code duplication and thereby also help reduce the amount of data to store and download.

Integrating BIOS and firmware updates into change-management frameworks

Existing Linux change-management frameworks can already handle Red Hat Package Manager (RPM) packages. By packaging BIOS and firmware updates and their related tools into RPM packages, the firmware-tools update functionality can be easily integrated into existing change-management frameworks.

Supporting additional operating systems and system types

The firmware-tools framework is written in Python, which is already included in Linux distributions. Using the firmware-tools framework for common tasks can help reduce the size of per-device-type inventory and executable applications. Dell anticipates that the inventory and executable applications will be compiled and built for each target Linux distribution version, and encourages these applications to be fully open source whenever possible to facilitate use in a wide variety of environments.

The firmware-tools framework also supports multiple target system types: servers, workstations, desktops, and notebooks. In many cases, administrators can use the same software to install BIOS payloads on any of these system types.

Splitting components into separate packages

The firmware-tools architecture splits the inventory, executable, and payload components into separate packages:

- **Inventory packages:** These include applications that report information about the current local system, including system vendor, system type, list of PCI devices, and so on. These packages use a plug-in scheme, allowing them to discover devices that do not fall under one of the standard categories—for example, an application that finds baseboard management controllers specified in System Management BIOS (SMBIOS) or Advanced Configuration and Power Interface (ACPI) tables rather than PCI.
- **Executable packages:** These include applications that install a given payload file into a specific device flash memory part. These packages also use a plug-in scheme, allowing them to easily integrate device-specific executables.
- **Payload packages:** These contain only the payload data files themselves (typically binary files) for a particular device.

The firmware-tools framework, like Dell Update Packages for Linux, can deliver package dependencies within the packages themselves and only requires administrators to download a single package per system or device type: firmware-tools uses the RPM provides/requires language to explicitly state inter-package dependencies, and uses an online distribution mechanism to resolve the dependencies automatically with a single download and install command. These capabilities allow multiple device update packages to use and share components such as the system inventory application, and help reduce code duplication.

If a new inventory application is released to add a feature or fix a problem, only that inventory application needs to be changed and pushed out to target systems, rather than the full self-contained package for each affected platform type. The framework also provides unifying applications such as inventory_firmware and apply_updates to implement runtime ordering of inventory, execution, and conflict resolution and notification for the plug-ins.

Using firmware-tools with change-management frameworks

This section discusses example firmware-tools usage scenarios for two change-management frameworks: Novell ZENworks Linux Management and yum.

Novell ZENworks Linux Management

Novell ZENworks Linux Management provides an easy way to keep Linux systems up-to-date. Novell releases packages containing new features and bug fixes in bundles and catalogs through a ZENworks Linux Management server; target systems check in with this server every few hours to download and install any updates. Firmware-tools places RPM packages containing firmware inventory, executable applications, and payload data into custom bundles on local ZENworks Linux Management servers, which target systems can then subscribe to.

To use firmware-tools with ZENworks Linux Management, administrators must do the following:

1. Obtain the firmware-tools and payload RPM packages and place them in custom bundles and catalogs on the ZENworks Linux Management server.
2. Subscribe the target systems to the custom catalogs.
3. Install the firmware-tools framework, vendor-specific tools, and payloads using the `rug` command.

The "Using firmware-tools with Novell ZENworks Linux Management" sidebar in this article provides example commands to carry out these steps; for additional details, visit the firmware-tools Web site at linux.dell.com/firmware-tools.

Yum

OpenSUSE, the Linux development project sponsored by Novell, uses a software update distribution tool similar to Novell ZENworks Linux Management. Yum (Yellowdog Updater, Modified) allows individuals to publish RPM packages on a Web server. Firmware-tools enables administrators to generate yum repositories containing the payload RPM packages, and systems can then subscribe to these repositories to obtain updated package versions. If yum is configured

USING FIRMWARE-TOOLS WITH NOVELL ZENWORKS LINUX MANAGEMENT

To populate a Novell ZENworks Linux Management server with firmware-tools packages and payloads and deploy those updates to target systems, administrators should first perform the following steps (as the root user) on the ZENworks Linux Management server:

1. Install ZENworks Linux Management on the server.
2. Create bundles using the following commands:

```
# zlman bundle-create bundle-dell-software
# zlman bundle-create bundle-dell-firmware
```

3. Obtain the RPM packages from the firmware-tools Web site.
4. Add these packages to the bundles, once per desired os-version-arch target:

```
# zlman bundle-add-package --installtype=upgrade
    bundle-dell-software sles-10-x86_64
    packagename.rpm
# zlman bundle-add-package --installtype=install
    bundle-dell-firmware sles-10-x86_64
    packagename.rpm
```

5. Create a catalog:

```
# zlman catalog-create catalog-firmware-tools
```

6. Add both bundles to the catalog:

```
# zlman catalog-add-bundle catalog-firmware-
    tools bundle-dell-software bundle-dell-
    firmware
```

Next, administrators (as the root user) should install the ZENworks Linux Management client on the target systems using the following commands:

```
# rug service-add https://zlmserver/
# rug subscribe catalog-firmware-tools
# rug install firmware-addon-dell
# rug install $(rug --terse what-provides
    "$(inventory_firmware -b)" | \
  awk -F \| '{print $3}')
# apply_updates
```

Note: Some versions of the `rug` command include a `solvedeps` function, which may be used instead of the second `rug install` line:

```
# rug solvedeps "$(inventory_firmware -b)"
```

to run automatically every night, it can download new inventory, executable, and payload components within hours of their release; administrators can also download the updates manually.

To use firmware-tools with a yum repository, administrators must do the following:

1. Obtain the firmware-tools and payload RPM packages and place them in a yum repository.
2. Subscribe the target systems to the yum repository.
3. Install the firmware-tools framework, vendor-specific tools, and payloads using the `yum install` command.

For more information about these steps, visit the firmware-tools Web site at linux.dell.com/firmware-tools.

Advancing the development of firmware-tools

Firmware-tools is designed to help significantly reduce the administrative burden of updating system BIOS and firmware versions on Linux-based systems. Its architecture has been demonstrated to work with Dell system BIOSs and Dell PowerEdge™ Expandable RAID Controller (PERC) firmware. A goal for the future is to expand the scope of products using this framework to include system types from other vendors, additional device types, and additional Linux and UNIX® OS types, along with additional packaging formats such as those used by Debian, Gentoo, and FreeBSD. Developers and system administrators alike are invited to join in this open source effort.

Matt Domsch is a Linux software architect at Dell and a community member of the Fedora Project Board. Matt has a B.S. in Computer Science and Engineering from the Massachusetts Institute of Technology and an M.S. in Computer Science from Vanderbilt University.

Michael E. Brown is a software developer and open source advocate on the Linux Engineering team at Dell. He was previously the lead developer for the Dell OpenManage™ Server Assistant application, where he led the switch to use Linux as the base OS for that product.

FOR MORE INFORMATION

Firmware-tools:
linux.dell.com/firmware-tools

Best Practices for Automated Enterprise Testing

A comprehensive deployment and verification strategy that includes test automation helps enterprises efficiently deploy and test software, hardware, and firmware by freeing administrators and testers from basic or repetitive tasks. This article discusses test automation fundamentals and best practices based on Dell enterprise product testing.

BY ERIC HOXWORTH AND MOHAMMED KHAN

Related Categories:

Application development

Software development

Testing

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

As the complexity of data center software and hardware stacks continues to grow, so does the complexity of ensuring that those stacks meet enterprise requirements. Effective testing and validation of data center environments typically requires automated tests in addition to manual ones. By minimizing human interaction during tests, automated testing helps efficiently verify that complex data center environments meet enterprise requirements.

Automated testing can provide two primary advantages: increased effectiveness and increased efficiency. It can increase effectiveness by allowing administrators and testers to complete more test scenarios than they could manually, enabling them to test deployment areas that would be impractical to test manually, and handling redundant or repetitive tasks, which allows them to focus on complex scenarios and helps reduce the chances of human error. It can increase efficiency by reducing test time, which can result in reduced time to service availability or a reduced need for human resources, and by enabling testing to continue beyond normal business hours.

Although automation is typically regarded as a cost-savings tool, the development and deployment of automated testing can reach a point of diminishing returns, where the costs of developing and deploying automation outweigh the benefits. For example, building a robot to physically insert and remove hard drives to simulate failures may be more costly than having a human tester perform the same task. Therefore, because of cost or time

Development difficulty	*Not possible: 0 points.* The test includes steps that cannot be replicated with a utility or other application.	*Hard: 1 point.* The test includes steps that three or more applications must perform.	*Intermediate: 2 points.* The test includes steps that two applications must perform.	*Easy: 3 points.* The test can be performed from a single browser or interface.
Time savings	N/A	*Low: 1 point.* The test takes less than 10 minutes to execute manually.	*Medium: 2 points.* The test takes more than 10 minutes but less than 60 minutes to execute manually.	*High: 3 points.* The test takes more than 60 minutes to execute manually.
Frequency of use	N/A	*Low: 1 point.* The test is not used regularly in any test cycles.	*Medium: 2 points.* The test is used regularly in at least one test cycle.	*High: 3 points.* The test is used regularly in multiple test cycles or in multiple configurations in at least one test cycle.

Figure 1. Sample evaluation criteria for determining which tests to automate

constraints, fully automated testing is not typically practical or desirable; enterprises should evaluate the appropriate ratio of automated to manual testing on a case-by-case basis.

Types of automation

Enterprises can take advantage of automation during multiple test phases in data center deployments, including installation, regression, stress, and load automation.

Installation. Installation automation enables administrators to efficiently deploy software, hardware, and firmware—including operating systems, patches, service packs, software updates, and security enhancements—and validate their successful installation, as well as launch installation programs. If administrators must deploy multiple systems with a similar configuration, they can use the same automation across these systems. Administrators should perform automated verification for the presence of files, accuracy of registry and configuration settings, and removal of temporary installation information, as well as the presence of specific hardware and the proper deployment of firmware.

Because completely automated testing is not typically desirable, developing a set of criteria to decide which tasks to automate is essential.

Regression. Regression automation enables testers to run specific tests to help ensure that changes to a system's software, hardware, or firmware have not affected the unchanged parts of that system. This testing should be developed so that it can easily and automatically run whenever a change is introduced into a system. For example, Dell performs automated regression tests on release candidates of Dell OpenManage™ software to help ensure that the changes to the software do not adversely affect the unchanged elements. In data center environments, a regression automation suite might check network connectivity after a software upgrade or database update functionality after a hardware upgrade.

Stress. Stress automation helps testers determine whether a system can withstand operating conditions at the boundaries of its capabilities, and includes running targeted applications to make a system operate at or beyond its operational limits. This type of automation can include validating disk subsystems to help ensure that they meet maximum throughput requirements.

Load. Load automation helps testers determine whether a system can handle a specified user workload, including Web loads, database loads, e-mail loads, and file I/O loads.

Automation criteria

Because completely automated testing is not typically desirable, developing a set of criteria to decide which tasks to automate is essential. These criteria should match enterprise needs, and can include such items as development difficulty, time savings, and frequency of use. These criteria should be weighted to account for enterprise priorities; for example, if meeting a schedule is the most important requirement, then time savings should be the highest-weighted criterion.

Figure 1 shows a simple set of criteria that an enterprise IT organization might use to determine which tests to automate. After evaluating each test based on these criteria and assigning the appropriate point values, the IT organization can generate automation scores using the formula *development difficulty* × *time savings* × *frequency of use*. Tests with the highest scores should be automated first, because they are the easiest to automate, provide the most time savings, and/or are used most frequently. A test with a score of 0 would not be automated.

Automation best practices

Automation software must be as robust as or more robust than the system or software being tested, so that testers do not confuse errors discovered by the automation with errors generated by the automation itself. To create a robust test automation suite, enterprises can implement the following best practices used by Dell in its own testing.

Automate tests only for stable releases. Creating automation for releases that are not static or stable can cause development churn and add to the cost of automation. In general, first releases or major system revisions are dynamic and typically not good candidates for automation. Enterprises can evaluate system release stability in the early stages of development and select automation candidates at that point.

> Measuring automation effectiveness helps ensure that enterprises are choosing the correct tests to automate and developing accurate automation.

Base automation on well-designed objectives. At Dell, test objectives are contained in test cases, which are housed in a test case library and used as the design for test automation. Each test case consists of a series of steps describing how to validate a particular feature of a deployment or product. Having well-designed objectives helps enterprises evaluate the usefulness of particular automated tests.

Treat automation development like software development. Automation development should follow a process similar to software development, including steps for planning, design and code reviews, user acceptance testing, deployment, and maintenance. A haphazard or ad hoc development process can result in increased costs and wasted resources to address missed schedules or unmet user expectations. It can also make maintenance difficult and inefficient, which, in addition to adding to resource costs, may cause the automation to fade into obsolescence as enterprise requirements change.

Create usability and maintenance standards. Frameworks for usability, reusability, and maintenance help facilitate standard development practices and enable rapid automation development. These frameworks define interfaces, control mechanisms, reporting functions, and common library functions that encourage reuse, simplify maintenance, and enable control over and visibility into test execution. They also enable a large body of developers to contribute to automation development in a consistent way.

Measure automation effectiveness. Measuring automation effectiveness helps ensure that enterprises are choosing the correct tests to automate and developing accurate automation. Example measurements include the number of times automated tests have run, the number of features automated testing covers, and the amount of time saved with automation.

Choose appropriate testing tools. Testing tools can be open source or fee based, utilize graphical user interfaces or command-line interfaces, vary in functionality and usability, and so on. Enterprises should select the most functional, cost-effective tool for a given task. They can use a prioritization matrix[1] and compare weighted requirements with features to help determine appropriate testing tools.

Efficient testing in enterprise data centers

Automation can increase the effectiveness and efficiency of enterprise testing, helping discover defects and demonstrate that systems meet enterprise requirements while minimizing human error and helping reduce test time—freeing enterprise IT organizations to focus on complex test scenarios. Following the automated deployment and testing best practices described in this article helps enterprises increase system uptime, meet time-to-market requirements, and enhance system and product quality.

Eric Hoxworth has 10 years of experience in software and hardware testing and currently leads the team at Dell that develops tests and automation for all Dell products. He has a B.S. in Electrical Engineering Technology from New Mexico State University.

Mohammed Khan has three years of experience in software and hardware testing and five years of experience in software and hardware development. He leads the team at Dell that develops test automation for Dell products. Mohammed has a B.S. in Computer Science from Florida International University.

[1] For an example prioritization matrix, see "Prioritization Matrix," by Syque, in *The Quality Toolbook*, syque.com/quality_tools/toolbook/Priority/priority.htm.

10 Gigabit Ethernet Unifying Fabric: Foundation for the Scalable Enterprise

10 Gigabit Ethernet is positioned to fulfill many expectations of Ethernet LAN technology that have been long anticipated but not yet realized by its predecessors and their alternatives. This article discusses the criteria for creating a virtual data center and how 10 Gigabit Ethernet can meet these criteria through its ability to function as a unifying fabric.

BY J. CRAIG LOWERY, PH.D.

Related Categories:

Data center technology

Data networking

Ethernet

Network fabric

Scalable enterprise

Virtual data center

Virtualization

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

The next generation of enterprise computing infrastructure is envisioned as a collection of standard, inexpensive, even disposable computing and storage components bound by specialized software into a distributed system. Often referred to as a virtual data center, these components in aggregate potentially offer unprecedented flexibility, enhanced reliability, and reduced total cost of ownership and systems management overhead.

Technical advancements since the virtual data center was described in *Dell Power Solutions* four years ago[1] have brought this goal within reach. Key to virtual data centers is the concept of a *unifying fabric,* also known as a converged network. Today, various classes of data center traffic use their own specialized interconnect fabrics, as shown on the left side of Figure 1: storage typically uses Fibre Channel for SCSI block-level access to storage area network (SAN) devices, high-performance computing cluster nodes use a high-performance interconnect such as InfiniBand to communicate with each other, and traditional networks typically use TCP/IP over Ethernet.

A unifying fabric, in contrast, enables all data center communications—including SAN, intra-cluster, and traditional network traffic—to use the same networking infrastructure and protocol set, as shown on the right side of Figure 1. All communication could take place through a single cable between each device and a concentrator or switch. In practical terms, various traffic classes are segregated for performance or security reasons, but they can still use the same networking technology and components. This simplification potentially leads to reduced equipment costs and a reduced number of support groups within an IT organization.

But what networking technology could serve as a foundation for building a unifying fabric? The requirements for this technology have been clearly understood for some time, and several candidates have been considered

[1] "Building the Virtual Data Center," by J. Craig Lowery, Ph.D., in *Dell Power Solutions,* February 2003, www.dell.com/content/topics/global.aspx/power/en/ps1q03_lowery.

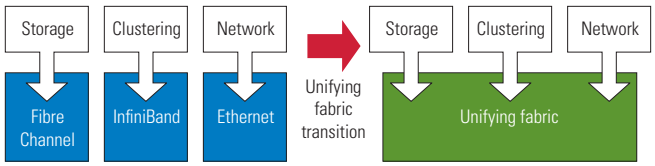


Figure 1. Traffic classes and the unifying fabric transition

and rejected. Most notable among these is InfiniBand, which has been and likely will continue to be successful in the high-performance computing inter-node fabric domain. Although it was specifically designed to be a unifying fabric,[2] it failed to gain acceptance as such because it was an unfamiliar, complex, expensive new technology, appearing well before its time with respect to the unifying fabric vision. And at that time, two key developments had yet to appear: virtualization and Internet SCSI (iSCSI).

Enabling convergence

Until recently, one of the primary inhibitors of unifying fabrics has been the tight coupling of software to hardware. In particular, an OS installed on a server contains drivers specific to that server's hardware components. Booting the OS depends on a series of events occurring in the BIOS and the OS bootstrap code that are configured during installation, after which the OS is bound to that server hardware; moving the OS to different hardware (for example, by removing the hard disk from one system and inserting it in another) may not work, even if the other system is identical to the original. Even after a successful OS boot, accessing disk volumes—especially those stored as RAID sets—depends on a mutually dependent software and hardware configuration.

This tight coupling largely nullifies the promise of unifying fabrics, which derive much of their value from putting compute and storage resources in different places on the fabric and connecting them dynamically as needed. Complementary technologies that break the dependency of software on hardware (notably operating systems and their hosted applications) are, therefore, a critical requirement for a successful unifying fabric.

Virtualization is an overloaded term in the current technical lexicon. However, when defined as a means by which software is dissociated from hardware, it becomes a key enabler for a unifying fabric. Server virtualization, such as that implemented by VMware® ESX Server and Microsoft® Virtual Server, clearly separates the software stack from the server hardware by creating a virtual machine (VM) environment in which the guest operating systems execute. It effectively makes every physical server appear to be identical to the VM guest operating systems, and therefore enables VMs to move freely to any server platform running the virtualization software.

Although server virtualization severs the ties between an OS and the underlying physical hardware, it does not provide the complete means for actually moving VMs between servers. Accomplishing this requires a type of storage virtualization in which VM state information resides on a remote storage device that can be connected to any physical hosting server participating in the fabric. The VM state information is contained in a virtual disk file—a large, flat file that mimics the functionality of a physical hard disk. To enable mobility, this file is kept on a remote storage device and accessed through the hosting virtualization software's file system facility. The VM can then be hosted by any physical server with access to the remote storage.

Enter iSCSI

In the early releases of ESX Server, mutual access to virtual disk files was accomplished by connecting all ESX Server systems to the same Fibre Channel SAN and multiple Ethernet networks, as shown on the left side of Figure 2. However, replacing the Fibre Channel SAN with an iSCSI SAN—the method of choice for implementing IP SANs—clears the way for a unifying fabric, as shown on the right side of Figure 2.

In addition to having a physical server host VMs that reside on an iSCSI SAN, a diskless server can boot from an iSCSI target logical unit (LUN). This method helps simplify boot-time logistics for servers because they only have to support one standard—iSCSI boot—instead of several potentially proprietary hard disk controllers (especially RAID controllers). Multiple diskless servers can iSCSI

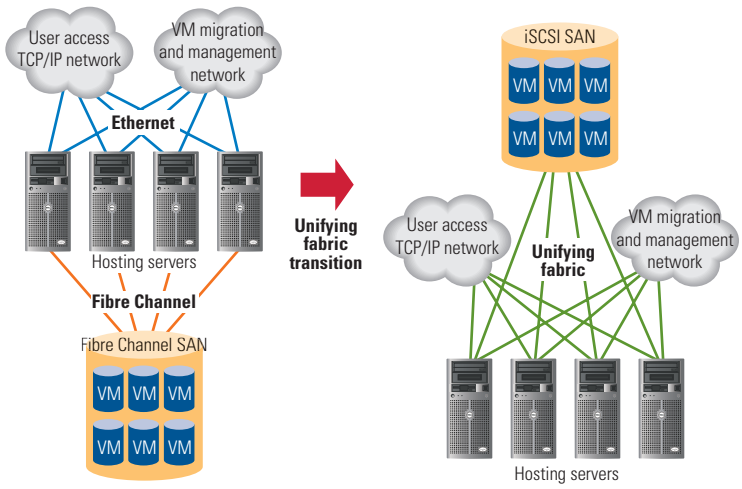


Figure 2. Virtual hosting infrastructure and the unifying fabric transition

[2] For more information, see "The Promise of Unified I/O Fabrics," by J. Craig Lowery, Ph.D., and David Schmidt, in *Dell Power Solutions*, February 2005, www.dell.com/downloads/global/power/ps1q05-20040191-Lowery.pdf.

boot a common virtualization software image, such as ESX Server, and then run multiple VMs—all across the same unifying fabric.

Creating a 10 Gigabit Ethernet unifying fabric

Figure 3 shows the conceptual building blocks of the Dell scalable enterprise based on a 10 Gigabit Ethernet unifying fabric. The Dell scalable enterprise seeks to standardize core data center elements to help simplify operations, improve resource utilization, and enable cost-effective scaling as needed. Virtualization is key to that vision, including VMs supported by server and storage virtualization technologies built on iSCSI and iSCSI booting, which in turn require TCP/IP. With its attendant optimization and offload technologies, 10 Gigabit Ethernet can serve as a unifying fabric foundation to help realize the scalable enterprise vision.

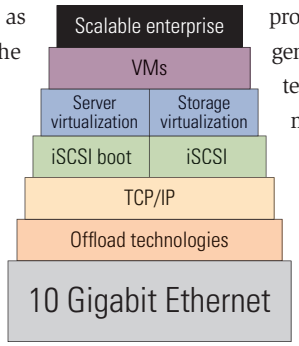


Figure 3. 10 Gigabit Ethernet as the foundation for the Dell scalable enterprise

Given that a unifying fabric must support storage, clustering, and traditional network traffic, and that server and storage virtualization technologies depend on iSCSI and TCP/IP protocols, a unifying fabric must meet the following requirements:

- Uses technology based on industry standards
- Becomes familiar to and trusted by the market
- Supports TCP/IP
- Coexists and interoperates transparently with existing interconnect technologies
- Provides sufficient performance (low latency and high throughput) to support iSCSI SANs and high-performance computing cluster interconnects
- Provides fault tolerance
- Offers cost-effectiveness

10 Gigabit Ethernet has emerged as the leading candidate to fulfill these requirements. As the next iteration of the established IEEE 802.3 Ethernet standards, 10 Gigabit Ethernet has a well-understood foundation on which to build; it also includes TCP/IP support and is designed to interoperate with previous generations of 10/100/1,000 Mbps Ethernet. SAN hardware from companies such as EMC provide both Fibre Channel and iSCSI capabilities, and legacy Fibre Channel enclosures can have a server front end that presents the Fibre Channel storage as iSCSI targets on the unifying fabric. Regardless, Ethernet gateways for Fibre Channel and InfiniBand already exist and can be further optimized for use with 10 Gigabit Ethernet.

Although Ethernet itself is not optimized for use as a unifying fabric, offloading technologies such as TCP/IP Offload Engine (TOE) and iSCSI offload help 10 Gigabit Ethernet provide sufficient performance for many application environments. Network interface card teaming, failover, and the robustness of the TCP/IP and iSCSI protocols can provide fault tolerance. As with previous generations of Ethernet, cost is likely to decrease as the technology is deployed in commodity volumes, eventually making it a ubiquitous and cost-effective technology.

Moving from vision to reality

As a unifying fabric candidate, 10 Gigabit Ethernet has received widespread positive support from the IT industry, and has been deployed since 2004 as a foundational networking technology with optical and heavy-duty copper-based interconnects for high-performance computing, where it has proven to be sufficient for a majority of inter-node communication tasks. PCI Express, protocol offloads, and multi-core processors can all contribute to removing bottlenecks that have previously prevented efficient use of 10 Gigabit Ethernet. A transition to 10 Gigabit Ethernet BaseT is on the horizon, at which time familiar CAT6 and CAT7 twisted-pair cable plants will likely be the cabling method of choice. 10 Gigabit Ethernet could become the true workhorse of enterprise computing, even supporting traffic types such as voice-over-IP telephony, videoconferencing, and remote computing.

Although it holds great promise, 10 Gigabit Ethernet as a unifying fabric requires much more than simply network adapters, switches, and cables. Management tools, methods, and best practices for iSCSI SANs and iSCSI boot (including boot image management) are needed. Refinement of protocol offloads, including integration into OS stacks and matching appropriate offloads to workloads, is still in the early stages of development. But even so, the consensus across the IT industry seems to be that a unifying fabric is a foregone conclusion, and that—with 10 Gigabit Ethernet identified as the key foundational technology—it is coming quickly. The transition may take several years, but 10 Gigabit Ethernet can take the industry one major step closer to realizing the virtual data center and the scalable enterprise.

J. Craig Lowery, Ph.D., is an advanced solutions technology strategist in the office of the CTO in the Dell Product Group. He currently focuses on strategic planning for enterprise solutions, particularly those that realize the Dell vision of the scalable enterprise. Craig has a B.S. in Computing Science and Mathematics from Mississippi College and an M.S. and Ph.D. in Computer Science from Vanderbilt University.

FOR MORE INFORMATION

Internet Engineering Task Force IP Storage Working Group:
www.ietf.org/html.charters/ips-charter.html

IEEE 802.3 Working Group:
grouper.ieee.org/groups/802/3

ENABLING BUSINESS AGILITY THROUGH VIRTUALIZATION

How an organization prepares for change often makes the difference between improving its competitive position or getting left behind.

Succeeding in today's business environment is all about agility. Organizations turn to technology to stay agile and proactive in the face of new challenges, but until recently, many technologies have not proven to be very agile themselves.

Why? In too many of today's data centers, applications are coupled tightly to specific operating systems, which are in turn tied to specific, often customized, hardware. Oftentimes, it is impossible for multiple operating systems to coexist on the same server, and for multiple applications to share the same operating system without conflict. Many companies thus find themselves burdened with thousands of servers, most running a single application. Adding server capacity requires long hardware acquisition lead times, and lengthy testing and certification processes, all of which hinder IT responsiveness. The alternative—deploying excess server capacity in advance—is a costly solution that risks obsolescence striking before that capacity is actually needed. Managing server sprawl is also expensive and resource intensive.

The lack of scalability and flexibility in traditional hardware configurations is why organizations are turning to server virtualization. Virtualization promises to make the data center more responsive to business needs. When virtualization is combined with the breakthrough performance, energy efficiency, and reliability of Intel's new Dual-Core Intel® Xeon® processor-based servers, it helps organizations scale without large increases to operating costs.

THE POWER OF VIRTUALIZATION

Virtualization transforms yesterday's rigid, complex infrastructure of individual servers, storage, and network hardware into a flexible virtual resource pool that IT can slice, dice, grow, and shrink dynamically to respond to new challenges, and to take a lead role on new business opportunities.

Virtual infrastructures are, fundamentally, based on virtual machines. Each virtual machine essentially is a container that encapsulates an operating system and associated application workload; it is completely isolated from neighboring virtual machines and independent from the underlying server, storage, and networking hardware. Multiple virtual machines can safely share the same server hardware platform, even if they are running different operating systems, and can be easily transferred across physical platforms regardless of hardware differences. IT administrators can add new hardware technologies without wrestling with compatibility issues or upgrading an entire software stack. Server hardware thus can be used to its full capacity, then easily added or re-provisioned to meet variable business demand.

Shared storage provides economies of scale for the virtual infrastructure by allowing scalable access to common storage arrays without constant hardware upgrades. Shared storage devices ease the creation, provisioning, and management of storage resources, enabling IT to more easily meet business needs via resource pools in a virtual infrastructure.

Enterprise system management technologies help streamline provisioning and management of the virtual infrastructure and underlying hardware components. Altiris Server Management Suite brings the management of both physical and virtual servers into a single console to help simplify and automate key functions including deployment, asset management, patch management, monitoring, and compliance. Advanced management capabilities offered through VMware Virtual Center—such as centralized management of physical and virtual resources, live virtual machine migration, dynamic data center resource optimization, and system backup facilities—gives IT the ability to optimize utilization of existing resources and improve overall operational efficiency. Resource re-provisioning to meet variable-demand loads can be addressed on-the-fly, and new IT services can be added in minutes instead of hours or days.

Companies like Jackson Walker LLP, for example, have been able to enhance IT scalability, manageability, and responsiveness dramatically by consolidating multiple physical servers into a few that are then divided into multiple independent virtual machines.

LESS TIME, WASTE, COST

"Recently, we had to create three new servers for our multi-tiered Ringtail Legal litigation support system," explains Steve McHargue, CIO for Jackson Walker, one of the largest law firms in Texas. "Thanks to our Dell and VMware virtualization solution, rather than having to spend several days procuring and provisioning three new pieces of server hardware, we were able to provision all of those servers in less than 30 minutes. We installed the application that afternoon, and it has worked marvelously."

Jackson Walker also consolidated its storage infrastructure using a storage-area network based on a Dell/EMC CX500 Fibre Channel array. "We had too many servers with huge direct-attached disk drives that were only 50 percent full," states McHargue. "We weren't

THREE STEPS TO THE AUTOMATED DATA CENTER

Virtualization is one of the three pillars that will take IT from yesterday's fragmented, brittle data center architecture to the efficient, highly automated data center of tomorrow. The other two pillars are management standards and an enterprise resource directory.

Management Standards: Today, every resource manager, such as a virtual machine monitor, has its own management console. As such, the adoption and implementation of new resources generally requires the adoption and implementation of a new management console. To achieve true flexibility, automation, and operating efficiency, tomorrow's data center will require management standards that allow resources to be managed in a consistent way from a common interface.

Achieving a truly automated, scalable data center will necessarily depend on standardization, rather than proprietary technologies, including adherence to evolving industry management standards such as the SMASH CLP standard for command-line interface scripts, and WS-Management and CIM-XML standards for accessing and exchanging management information across the enterprise.

Standards such as these also provide interface stability. Key to ensuring this stability over time is to keep the interface between the workload and physical computing resources stable, even as elements on either side of the interface change.

Enterprise Resource Directory: With virtualization and common management standards in place, the next step is to create a single comprehensive enterprise resource directory for all rules, relationships, and resources, instead of the fragmented pockets of knowledge found across dozens of products and management tools in today's data center. A single directory will be the key that enables dynamic, automated, policy-based resource allocation to become a reality.

getting very good disk utilization by having a lot of servers with dedicated storage that wasn't networked. With our Dell/EMC SAN in place, there is far less waste and a lot less investment required to procure the amount of disk space we require. And we know that if an application needs more storage over time, we can easily scale to meet that need."

INFINITELY SCALABLE

Virtual infrastructures can dramatically increase IT agility by enabling IT to cost-effectively scale its computing, storage, and networking resources to support dynamic business requirements.

Virtual infrastructure makes it more practical to build advanced, "scaled-out" IT infrastructure by adding applications, standards-based servers—such as Dell PowerEdge servers powered by Dual-Core Intel® Xeon® processors—and storage quickly and incrementally as they're needed. This avoids costly "scale-up" strategies based on proprietary and expensive, CPU-dense SMP servers that are often over-provisioned to allow room for future growth. A scale-up strategy can be more costly and inefficient because it requires organizations to overspend precious IT budgets today in anticipation of computing capacity needs tomorrow, and often locks them into technology that becomes obsolete before it is needed. Scaling out enables a much more efficient pay-as-you-grow strategy that helps deliver sustained, incremental performance improvements at consistent, competitive prices.

The other, often overlooked, attribute of virtualization that contributes significantly to business agility is the power of encapsulating complex server/software configurations into a file using VMware software to create a virtual appliance.

Virtual appliances are preconfigured software applications, such as a database server or a network firewall, encapsulated in a virtual machine that can be run on any server that is running VMware's virtualization software. It is not a stretch to say that this approach might fundamentally change the way we think about computer configuration and software distribution. Here's why: organizations, when purchasing software, often spend more on professional services than they do on the software itself because of the complexity of configuring server software (OS, middleware, applications), and then replicating and "transferring" that workload from one server to another. With virtualization, all the complexity of a server can be captured into a file that will run on any server that's running the associated virtualization software.

Deploying new services with virtual machines takes minutes instead of hours or days that are typically spent installing and configuring an operating system and application software.

PARTNERS FOR PROGRESS

Dell's commitment to standards-based virtualization architectures and long-standing relationships with industry leaders Altiris, EMC, Intel, and VMware offer customers integrated, scalable, and reliable virtualization solutions that are best suited for their individual needs. By bringing together the key components—servers, virtualization infrastructure software, management, and storage—these partners deliver a tightly integrated solution that helps address top IT challenges.

By offering a single source for proven, market-leading virtual infrastructure and comprehensive services to assess, design, implement, and support, Dell reduces the complexity of implementing a virtual environment. Customers can rest assured that they are on solid ground—and backed by trusted partners—when taking the first essential steps to tomorrow's virtual data center today.

For more information, please visit **www.virtualization.ziffdavis.com**. ■

POWER TO THE PEOPLE!

Team the power of Dell hardware with software from Microsoft for people-ready business. Then watch productivity soar.

Give people the power to be more productive by making Dell your single source for business-building IT solutions that include Microsoft® Windows® Vista,™ Exchange Server 2007 and Office 2007. Then watch people work together. Easily. Efficiently. Securely.

- Dell engineers hardware for Microsoft software.
- Dell customizes solutions to handle new challenges.
- Dell delivers technical support directly to you.

There are lots of ways Windows Vista, Exchange Server 2007 and Office 2007 can power your enterprise. But there is one smart way to make it happen. By making Dell your one-stop resource for technology that's optimized and customized for you.

And for the expertise to turn your productivity wish list into a competitive reality. **Power to the People!**

Microsoft®

DELL Purely You
www.dell.com