

HIGHLIGHTS

- » Easy Server/Storage Monitoring
- » NetApp HW Platform Upgrades
- » Better Backups for VMware
- » Best Practices: x64 and iSCSI
- » Compute Clusters Unleashed

Examining the Future: Trends in Storage Technology

Dave Hitz, Founder and EVP, NetApp

A look at how storage has evolved over the past 20 years and what to expect in the next 20.

[More »](#)



BLOGGING WITH DAVE

Dave Hitz, NetApp Founder and EVP
Winner: Storage Haiku Contest

So many users
Not enough disks in array
Thin provisioning



- » Dave's Blog
- » 2006 Storage Haiku Winners

DRILL DOWN

- » **New Industry Performance Record**
NetApp achieves first-ever NFS benchmark result of greater than 1 million operations per second.
- » **Data ONTAP GX Launch**
Unleash the value of compute clusters with flexible storage building blocks.



ADMIN RESOURCES

- » **'Free' Featured Tool: Cacti**
Graph usage and performance for hundreds of servers/systems/devices.
- » **Implications of the Microsoft® TPC Chimney Announcement**

TIPS FROM THE TRENCHES

Scalable Reliability for SQL Server with iSCSI and Windows® x64

James H. Lunsford, Jr., Applications Development Manager,
Bolt, Cummings, Conners & Berry, PLC

A comparison of x64 performance in Fibre Channel and iSCSI environments plus best practices involving round robin sessions, redundancy, LUN configuration, and more.

[More »](#)

Love VMware but Hate Backups?

Vaughn Stewart, Systems Engineer, NetApp

A discussion on block- and file-level options for backing up virtual machines plus tips for improving backups without changing storage vendors. (fyi: VMware ESX Server 3.0 will provide native iSCSI and NAS support!)

[More »](#)

ENGINEERING TALK

What Makes the FAS6070 So Fast?

Howard Gage, Manager, HW Engineering
Tom Holland, Sr. Engineer, Platform HW Engineering
Naresh Patel, Sr. Engineer, Performance Engineering

NetApp engineers discuss architectural changes that enabled a 100% performance improvement from the previous platform.



Includes comments on the transition to a native 64-bit architecture, HyperTransport™ links, PCI Express, and increased NVRAM memory.

[More »](#)



DAVE HITZ

Founder and Executive Vice President, NetApp

As executive vice president, Dave is responsible for the future strategy and direction of Network Appliance. He and James Lau initially founded Network Appliance in 1992 with a desire to simplify storage the way Cisco simplified networking. Prior to NetApp, Hitz worked as a senior engineer at Auspex Corporation and in engineering roles at MIPS Computer. Before his career in the computer industry, Hitz worked as a cowboy, where he got valuable management experience by herding and branding cattle.

Examining the Future: Trends in Storage Technology Adoption

By Dave Hitz

Network Appliance spends a lot of time looking at storage trends and thinking about how emerging technologies will change the industry. Today I'm feeling ambitious, so I'm going to share a quick overview of what we believe the next 10 to 20 years of storage are going to look like.

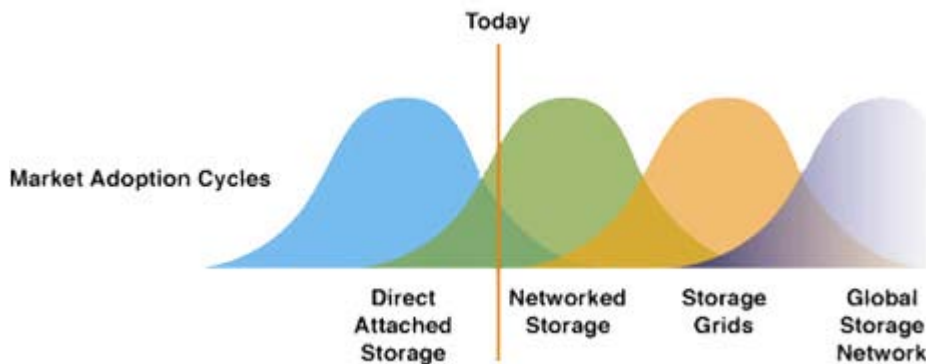
Like any technology, storage adoption generally moves in waves. New technologies emerge in response to the limitations and challenges of existing technology, while other technologies mature and then move into decline.

The NetApp model representing the evolution of storage solutions technology includes four major waves:

- Direct-attached storage
- Networked storage
- Storage grids and scale-out architectures
- Global Storage Networks

Wave 1: Direct-Attached Storage

Until the early 1990s, the primary storage model involved individual disks attached directly to individual servers. Although the disks themselves are affordable, associated costs—licensed server connections and management—are high relative to more recent technologies, and management can be particularly challenging.



Wave 2: Networked Storage

Networked storage, the second wave of innovation, was developed in response to the

RELATED INFORMATION

- [Exclusive Early Access: Examining the Future \(pdf\)](#)
- [From Reactive to Proactive Admin](#)
- [Scale-Out Architectures \(Podcast\)](#)
- [Data ONTAP GX Launch Site](#)
- [NetApp Shatters NFS Performance Record](#)

Compute Clusters Unleashed: Technical Toolkit

If you run data-intensive high-performance computing applications, the newly released Data ONTAP GX storage systems can help you get the very most out of your Linux® compute clusters.

Data ONTAP GX provides flexible storage building blocks for designing and deploying scalable storage environments.

Key benefits include:

- Breakthrough performance with true linear scalability
- Management simplicity and the ability to scale to multiple petabytes under a single global namespace
- The ability to transparently move data between controllers and storage tiers
- Advanced resiliency and recovery capabilities

Learn how your business can benefit from unleashing the value of your Linux compute clusters with the most practical high-performance scale-out systems available today:

limitations of using direct-attached storage. With the move to networked storage, capabilities that used to be higher up the stack and outside the storage system have moved into it. For example, RAID and volume management are both add-on capabilities in the DAS world but are built-in capabilities of current networked storage systems.

The bulk of this market is still ahead of us, which means there is a big opportunity for innovation. We think the innovation is going to come in two spaces. One of them involves managing lots of disk drives, and the other involves managing and automating the interface to these storage systems.

Why is dealing with disk drives so important? If you think about direct-attached storage, you've got one server, and you've got some storage attached to it. You only have as many disks in that box as the server needs. Once you go to network storage you can have as many servers as you want talking to the same box of disks, and so you get absolutely enormous systems with hundreds of disks.

You just can't even think about dealing with that many disks one at a time. You have to go with some kind of virtualization model. That leads to what's called a scale-up architecture. You start with a small box that has a smaller number of disks, then bigger boxes, and then at the high end of your product line the biggest box you can imagine, because that lets you have fewer boxes to manage.

Wave 3: Storage Grids and Scale-Out Architectures

Even with the largest systems you can imagine, in large environments people still end up with dozens or hundreds of separate systems. In the same way we can virtualize disks within a single box so that it feels like a big pool, we need to figure out how to virtualize entire systems.

That's what scale-out is all about, and that's why innovation in storage grids based on scale-out architecture represents the next leap in storage technology.

A scaled-out storage system architecture breaks the system boundary by incorporating multiple systems into a common pool of storage that can be managed and accessed as if it were a single system. So, yes, you've got multiple nodes and you keep adding more boxes—just like you add more disks in a big storage system. But you don't think of them as lots of separate boxes; they work like one big "system."

Initial efforts in this zone focus on supporting large, compute-intensive environments that would overwhelm the throughput of any single machine. Administering storage when you are supporting thousands of nodes is daunting without the ability to consolidate the management of all the storage required. By consolidating the management of storage systems, scale-out technology further simplifies storage and data management.

In the long term, the scale-out model will be crucial to the ongoing evolution of storage technology because of the operational efficiencies that the model enables. Because of the many benefits of scale-out architectures, numerous vendors will offer scale-out technologies for storage and data management.

We expect that grids from different vendors will interoperate only to a limited extent and that the most sophisticated features from each vendor will only work within their own grids. Islands of data will become huge, veritable Australias of data. For robust interoperability across all Australias of data, we need to look to the next wave.



Islands of data will become huge, veritable Australias of data.

Wave 4: Global Storage Networks

The ultimate wave in our model represents the Global Storage Network. At this point universal standards allow boxes from all vendors to work together seamlessly. They also enable automation, so all of the data management tasks that chew up untold hours of time go away. The entire storage network is managed as a single entity. It is data utopia.

Before you get too excited about this final wave, take a look at the time from the beginning of that second wave—direct-attached storage—going back into direct-attached storage today.

- [Intro to Data ONTAP GX: NetApp Scale-Out Storage Concepts and Benefits](#) (30-page technical report)
- [Data ONTAP GX tech specs](#)
- [NetApp Shatters NFS Performance Record](#)
- [Data ONTAP GX launch site](#)

Best Practices for Data Center Design

Dave Hitz and other experts share their perspectives on achieving:

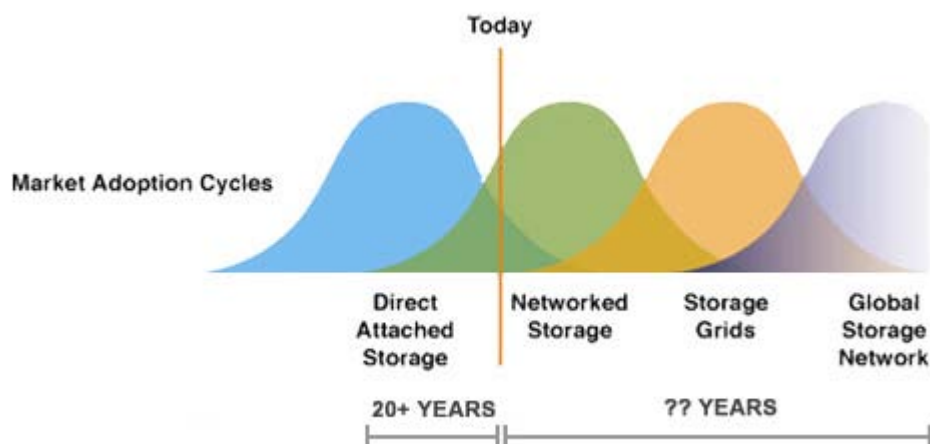
- Maximum flexibility
- Consistent application service levels
- Massive scalability
- Highest application availability
- Highest staff/asset productivity

Watch the [TechTalk Webcast](#). Sign up for the [Proven Thursdays Webcast Series](#).

The Impact of Scale-Out Architectures on HPC



Watch Tom Georgens, EVP and GM of NetApp Enterprise Storage Systems, discuss scale-out technology at the [Data ONTAP GX launch site](#).



Keep in mind that network storage started in the mid 1980s. That's a 20-year timeframe from the beginning of that second wave to today. I'm not going to argue that the time scale stays constant going forward, but we're talking about long, slow transitions.

This last wave depends on industry acceptance of universal storage and data management standards that aren't even close to existing today. The industry is only now starting to experiment with the data management models—such as scale-out architectures—which will give rise to these standards.

Given how long it has taken the industry to move from the beginnings of networked storage to where we are today, we don't expect that the Global Storage Network will be a focus of innovation until many years to come.

So What Do These Trends Mean for NetApp?

The network storage market is still where a lot of the opportunity is today, so you'll see NetApp continuing to invest in making it easier to manage a lot of disk drives. In May, for example, we introduced the new FAS6000 platform, which can scale up to 500TB and support tiered, consolidated storage for hundreds of applications on a single system.

Ultimately, however, we believe that the vendors that are working to be innovative with storage grids and the scale-out model will be the leaders of the next decade. This is why the new Data ONTAP® GX storage systems are incredibly significant.

Data ONTAP GX eliminates the single system barrier and enables many controllers to be combined so they appear to applications as a single entity and scale as a single entity. In addition, data can be moved between different controllers and/or tiers of storage without disrupting applications or users.

Initially Data ONTAP GX will focus on meeting the demands of high-performance computing customers, but this release is just the beginning. For high-performance computing environments we'll scale to even higher levels of performance and capacity, while over time we'll fold in additional capabilities to bring the benefits of scale-out to the full range of customer environments.

To help explain how all of these pieces fit in a technology sense, the NetApp executive team has created a [white paper](#) outlining our vision for the future of storage and how we're investing for virtualization, scale-out infrastructures, and heterogeneous data management.

You're invited to read an advance copy, and if you have any comments I'd be interested in hearing them.



Read [Examining the Future of Data Management](#).

Comment on this article



JAMES H. LUNSFORD, JR.

Applications Development Manager, Boulton, Cummings, Conners & Berry, PLC

Since discovering the joys of programming with a TI-994a at the age of 11, James has become a mostly self-taught expert in systems and applications. In 1990, James joined Boulton, Cummings, Conners and Berry (BCCB), a Nashville-based law firm. Today, James is responsible for custom application development, customization of off-the-shelf applications, system architecture, and integration. He has been working with NetApp storage for more than a year. When James is not in the data center, he can be found serving as pastor to a small rural parish outside Nashville, Tennessee.

Scalable Reliability for SQL Server with iSCSI and Windows x64

By James H. Lunsford, Jr.

For certain applications, the 32-bit address space of the x86 architecture and standard Windows is a significant limitation. The total addressable memory is limited to 4GB (no single, widely accepted standard exists for extending memory), and a single process is seldom able to use more than 2GB of memory. This is a particular problem for databases, where the ability to cache large amounts of data in memory can significantly boost performance. The transition to a 64-bit architecture is complicated, however, by the need to continue running existing 32-bit applications.

The release of x86-compatible processors with 64-bit extensions (x64) combined with Windows Server 2003 x64 Edition delivers the benefits of a 64-bit address space without sacrificing compatibility with 32-bit applications. Because every 32-bit process can be allocated up to 4GB of memory, x64 offers substantial performance improvements for databases and other memory-intensive applications without disruptive application changes.

When it came time to upgrade our Microsoft® SQL Server infrastructure last year, our team felt that x64 was the way to go because of the then imminent release of 64-bit SQL Server 2005. We wanted an architecture that would best fit with the direction we knew we would be taking in the future. Early internal testing with x64 demonstrated significant advantages in data throughput, even for our current 32-bit applications.

A Comparison of x64 Performance in FC SAN and IP SAN Environments

We initially planned to use a Fibre Channel SAN with the new server environment, as we had done previously. However, we learned that a vendor-specific module (VSM) is required for multipath I/O (MPIO), which allows a SAN to use multiple network connections for greater performance and reliability. This capability is built into the Microsoft iSCSI initiator for IP SAN. After the discovery that this module was not yet available from NetApp for x64 and Fibre Channel (a NetApp module was made available in early 2006), we realized that our only option was iSCSI. We were skeptical, but willing to give iSCSI a chance if it was capable of meeting our requirements.

We did everything possible to enable an apples-to-apples comparison. All performance tests used the free Microsoft [SQLIO Disk Subsystem Benchmark](#) available on the Microsoft Web site. The same server and storage system were used for both the Fibre Channel and iSCSI tests. One significant difference was that MPIO was engaged on the iSCSI test and was not engaged on the FC SAN test. The other thing that differed was the network between server and storage, Fibre Channel in one case, Gigabit Ethernet (GbE) in the other. In both cases, total available network bandwidth was 4Gb per second.

In the following table, I've used the performance of the original configuration (32-bit

RELATED INFORMATION

- [Best Practices: SQL Server on NetApp Storage \(pdf\)](#)
- [Storage Blog by Nick Triantos, NetApp SAN Systems Engineer](#)
- [Reducing Time-to-Deployment: Improving Agility for Test/Dev](#)
- [SQL Server 2005 Performance Comparison: FCP vs. iSCSI \(pdf\)](#)

SQL Server 2005 Performance in SAN Environments

NetApp recently teamed with Microsoft to study the performance of 64-bit Microsoft SQL Server 2005 on Fibre Channel and iSCSI. (These tests were performed on systems using Intel Itanium processors, not the x64 processors described in this article.)

Key conclusions:

- iSCSI is an equally viable solution for large-scale database workloads.
- Performance was inhibited in each case only by available server CPU.
- Both protocols performed at 78% of total available bandwidth.

For complete details on the test environment and results, read [the white paper](#).

server with 1Gbps FC) as a baseline and reported the results of each SAN relative to those baseline numbers. The results speak for themselves.

Test Configuration	IOPS*	Data Throughput
Original (32-bit) baseline		
FC w/out MPIO (2-2Gbps, x64)	2.7X	5.9X
iSCSI w/MPIO (4-1Gbps, x64)	5.2X	5.2X
[*]IOPS=I/O Operations Per Second		

Editor's note: These results are based on a synthetic workload generator and should not be understood as representative of the performance of any current or future Microsoft or NetApp product.

The IP SAN delivered the best transactional performance we have measured to date and comparable data throughput on our key benchmarks. Subsequent testing with SQL Server demonstrated that x64 offered at least twice the data throughput on real-world applications. Even our 32-bit applications benefited greatly in the x64 environment. At this point we were sold on x64 and impressed by iSCSI, so we chose and deployed an IP SAN for our new SQL Server 2005 environment.

During a [NetApp Webcast](#) earlier this year, I mentioned the extensive testing we'd done while optimizing our SQL Server environment. After the session, Tech OnTap asked me to share our best practices for x64 and iSCSI on production servers. Following are the three main practices that our team uses to ensure performance, reliability, and scalability.

James's Best Practice #1: Use Separate Networks and Round Robin Sessions

To achieve the best performance, we dedicate four GbE NICs on each server and four interfaces on each storage system for iSCSI, just as we did in the test environment. We use virtual LANs (VLANs) on our Cisco 6000 series switches to isolate iSCSI traffic from user traffic for better performance and data security. (Separate switches dedicated to iSCSI would be a viable alternative.)

Rather than using EtherChannel to trunk the network connections together, we allocate each NIC a separate subnet and use the MPIO and round robin capabilities of the iSCSI initiator to spread requests across the four network channels. I attribute the great transactional performance we measured with iSCSI to this ability. In contrast, adding more FC connections didn't improve benchmark results. In addition, a dedicated IP switch and dedicated NICs end up being much cheaper than comparable Fibre Channel infrastructure. I call this configuration N+3, scalable reliability because I can lose three NICs and still have access to storage.

James's Best Practice #2: N+1 Redundancy or Greater at All Levels

We extend this reliability throughout our server infrastructure to achieve N+1 reliability or better throughout. We have two, 2-node server clusters that use Dell 6850 servers (x64 capable), each with 16GB of RAM and six GbE ports (four for iSCSI, two for regular network traffic). Each cluster is configured in an active/passive configuration to protect against server failure, and SQL Server applications are load balanced between the two clusters. We use Microsoft Virtual Server on each passive server to create an N-tier farm that replaces some older blade servers, so the passive nodes aren't just sitting there waiting for the active node to fail.

On our Cisco switches, we make sure that the network connections from servers and storage are striped across multiple blades and multiple switches, to ensure redundancy at the switch level as well as at the NIC level.

Our NetApp FAS3020c storage system is an active/active, two-head cluster that gives us N+1 protection against hardware failures. We use NetApp RAID-DP™, which protects against up to two simultaneous disk failures in a single RAID group. We also allocate two hot spare drives for each aggregate, so that any failed drive is immediately rebuilt to limit the risk from multiple drive failures. (To learn more about NetApp hardware resiliency features, check out [The Private Lives of Disk Drives](#).)

Environment

- **Server:** Dell 6850 with 16GB of RAM
- **Operating system:** Windows Server 2003 x64
- **FC Infrastructure:** Server with dual-port FC HBAs, Brocade 3250 eight-port switch—2 FC ports on 3020c (all 2Gbps)
- **IP Infrastructure:** Server with four GbE NICs—Cisco 6000 series switch—two dual-port NICs on NetApp 3020c, Microsoft iSCSI 2.0 Initiator
- **Storage System:** NetApp FAS3020c



Transitioning to IP SAN

In a recent NetApp TechTalk Webcast, James explained the process he went through when he transitioned his SQL Server and Exchange environments to a NetApp SAN.

He also highlights results: the full migration was completed in just 12 hours!

View the [Webcast](#).

NetApp Best Practice Tip for SQL Server Backups

When in doubt, NetApp generally recommends backing up each database every hour and transaction logs every 30 minutes during production time.

Learn more [SQL Server Best Practices \(pdf\)](#).

Key Questions for Choosing Between FCP and iSCSI

1. Does your application require block or file I/O?
2. What levels of redundancy do your applications require?
3. Do your applications use small or large block I/O, how many IOPS are being pushed, and how much bandwidth do you need?
4. Is your current infrastructure pure Ethernet, or do you have Fibre Channel fabric?
5. How do requirements differ between your main data center, regional data centers, DR facilities, and remote offices?
6. Is getting the "lowest cost" solution a top priority?

Find out more. Read [Mapping Out a Strategy to Minimize DAS](#).

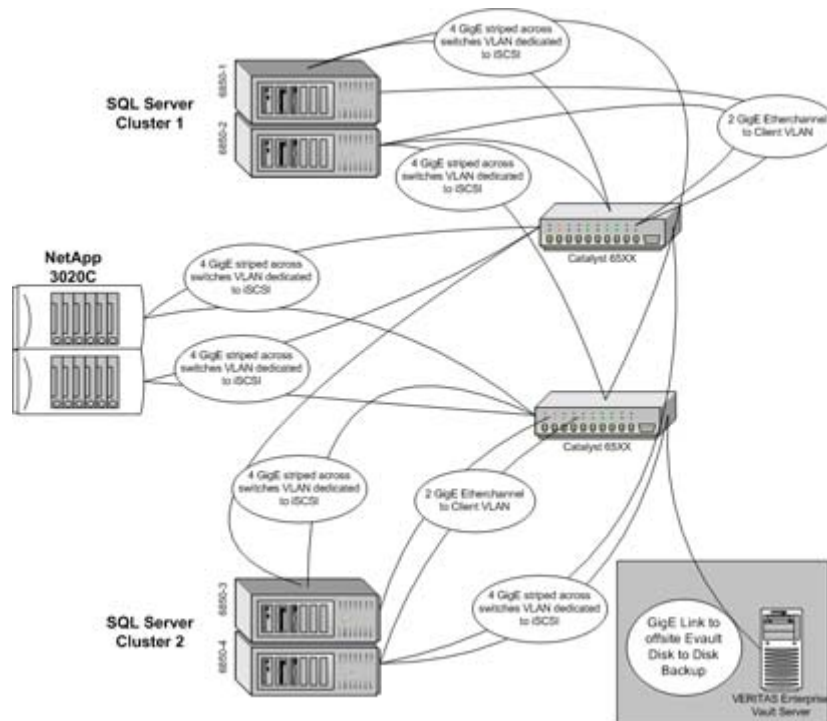
James's Best Practice #3: Configure Multiple LUNs and Spread the Load Across Storage Systems

To achieve optimal storage performance, every SQL Server database is configured with a minimum of two LUNs: one for data files and one for transaction logs. For bigger databases—over the 20GB mark—we split the data files into file groups and spread them across multiple, separate LUNs. This gives us maximum benefit from the round robin scheme.

I also like to split the load between the two storage systems. For instance, my busy accounting database is on the first storage system, and my document management system is on the second. The SQL Server system tables are also on a separate LUN. In my own personal experience, using a lot of separate, small LUNs ensures maximum performance for each application.

The Result: A Robust x64 Server Architecture

A picture is worth a thousand words:



[Click for larger view](#)

This is our SQL Server infrastructure today. Before the upgrade, our backup process took three hours, and we couldn't do routine verifies because they took too long. Today we dump to a separate LUN on the NetApp 3020 and verify it in only 12 minutes. We're very pleased with the results, and in my opinion there is no question that both x64 and iSCSI are well suited for deployment with SQL Server, Microsoft Exchange, and other Windows applications.

Editor's note: This article describes strategies and methods developed by a NetApp customer. Our legal department requires that we point out that this article may include workarounds and suggestions not documented in NetApp best practices materials and not supported by NetApp.

Comment on this article



VAUGHN STEWART

Systems Engineer and Resident VMware Expert, NetApp

Vaughn has been a MSSP consultant and systems administrator architecting and supporting a wide range of solutions including VMware virtualization software. Vaughn is currently certified on solutions offered by Microsoft, Cisco, Sun Microsystems, IBM, and Network Appliance. Vaughn joined NetApp in 2000 and during the past six years he's worked with hundreds of customers and assisted directly with numerous VMware implementations and joint solutions.

Love VMware but Hate Backups?

Virtual infrastructures provide a simple solution to the challenges associated with distributed server architectures. More than 20,000 enterprises worldwide use server virtualization software, and the launch of [VMware Infrastructure 3](#) with native iSCSI SAN and NAS support makes this type of consolidation a viable option for companies that hadn't previously considered it.

IT teams that deploy virtual infrastructures usually report that server virtualization makes their lives easier overnight by lowering the total amount of physical servers, network ports, floor space, maintenance contracts, and electricity required to run their data center operations. Once the honeymoon is over, however, many of these same IT teams find themselves trapped in a backup nightmare.

Whether you're considering an initial foray into server virtualization or are already feeling trapped in a love/hate relationship, backup and recovery is a critical consideration. To answer some of the top questions posed by NetApp customers, Tech OnTap turned to Vaughn Stewart. Vaughn coauthored the new technical report [NetApp and VMware ESX Server 2.5.x: Building a Virtual Infrastructure from Server to Storage \(pdf\)](#) with NetApp Professional Services Consultant Mike Slisinger.

Q. I've consolidated 10 physical servers onto a single VMware ESX server. At first it was amazing: we cut hardware and increased CPU utilization, and not a single end user noticed the difference. But now the fact that a single service disruption could impact my entire environment is keeping me awake at night. Do I have to live with this level of insecurity?

Vaughn: Unfortunately your concerns are well founded: with any consolidation effort you are in essence placing more eggs in your proverbial basket, so the redundancy and resiliency of your consolidation platforms are critical decisions in your design. The first step to upgrading your virtual infrastructure platform involves switching from local to shared (or networked) storage.

Here are a couple benefits of networked storage for consolidated environments:

1. **Drive failure protection.** It's a sad fact of life that drives fail. To get around this, many companies move to a form of data protection that is more resilient than the RAID5 most physical servers are deployed with. There are three key options: [RAID10](#), [RAID6](#), and [RAID-DP™](#).

Which RAID option is right for you? All three protect against double disk failures without impacting the availability of your virtual infrastructure. Unfortunately, RAID10 requires 100% storage overhead (which doubles your storage costs), and RAID6 suffers from significant performance overhead. The most efficient option is RAID-DP technology, which offers protection from simultaneous double drive failures while providing the high performance of RAID10 and the cost savings of RAID6. (See a related Tech OnTap article about [disk drive resiliency](#).)

RELATED INFORMATION

- [VMware Raises the Bar with VMware Infrastructure 3](#)
- [NetApp and VMware ESX Server: Going Virtual from Server to Storage \(pdf\)](#)
- [Using a NetApp SAN with VMware for Storage/Server Consolidation \(pdf\)](#)
- [Fueling the Enterprise Grid with Whole Application Virtualization](#)

Information-Optimized Backups

It used to be that a larger library, faster tape drive, upgraded network, or new backup server kept backups running.

Over the past five years, however, disk capacities have grown 20x, with a scant fourfold performance improvement for backup. Today, many companies are turning to smarter backup solutions that:

- Reduce backup storage consumption by eliminating redundancy within backup files
- Shrink backup windows by copying only incremental data changes
- Ensure reliable, recoverable backups by validating backups as they are created

Read [When Brute Force Isn't Enough: The Value of Information for Backups](#).

A Storage Networking Appliance

2. **Consolidated backups.** Since networked storage enables multiserver access, it makes your life significantly easier when it comes to backups and workload management.

The bottom line is that shared storage will let you build in greater protection and take advantage of more VMware features and management tools. Combining VMware VirtualCenter management software with shared (or networked) storage gives you the ability to dynamically move workloads around without service disruption.

By the way, if you're not currently using VirtualCenter, stop reading this article now and download it right away. VirtualCenter is an incredibly useful tool for managing multiple virtual machines, and, if you're a registered VMware customer, a 30-day license is 100% absolutely free. I guarantee you won't want to give it up.

Q. I've been using VMware for 12 months with about 50 virtual server machines. We've got less hardware, and my backup processes haven't changed, but suddenly I'm spending almost all my time managing backups. How do I stop the madness?

Vaughn: Many customers find that their consolidations are so successful they outgrow the backup bandwidth of their virtual infrastructure. Consolidating tens of physical servers onto a single server often creates a disproportionate ratio of data to physical server bandwidth. While not all VMware deployments will initially face backup challenges, it has quickly become a hot-button issue in the virtualization community. Check out conference agendas from VMworld 2005 and note the number of times the word backup is on the agenda.

Unless you want to permanently devote your nights and weekends to backup processes, you're going to have to make some changes to how you approach backups.

Today, users of VMware's ESX Server have two options for backing up virtual machines: backing up virtual storage either as physical storage or as files. Each alternative solves the problem of protracted backups, but both options have drawbacks when it comes to recovery. Some customers are waiting for the release of the new VMware Consolidated Backup product which is part of VMware Infrastructure 3, but the new software suite was not yet available when this article was published. NetApp technology can make any of the existing solutions work today and, if you want to start doing consolidated backups right now, you can.

Q. I'd love to move to disk-based backup, but it looks too rich for my budget. One vendor tells me I'd need as much as 100% additional storage for each backup.

Vaughn: Those vendors are doing the best they can with their traditional disk-to-disk (D2D) backup solutions. If you are backing up the files that make up the VM (the virtual disks and configuration files), most vendors can only offer two choices: (1) a "mirror copy" approach, which offers optimal performance but requires a 100% additional storage requirement for every backup; or (2) a "copy out" snapshot option, which only copies changed data blocks but comes at such a performance penalty that you'll be hard pressed to find any of their customers actually doing it.

The inherent negative features of traditional disk-based backups do not apply in NetApp storage environments. The unique NetApp Snapshot™ technology enables you to make instantaneous, full backups without performance penalties and at a cost of only a fraction of the original disk space. When combining NetApp Snapshot copies with VMware ESX server, customers are able to back up their entire virtual infrastructure within seconds. (Stephen Manley explains NetApp backup efficiency at length in his Tech OnTap article about smarter backups.)

NetApp Snapshot technology opens up a number of other possibilities. Snapshot copies can be backed up to tape and/or replicated to another facility, VMs can be restored, individual files may be recovered, and clones can be instantly provisioned for test and development environments. You can also utilize this solution with VMware SAN LUNs (which are also referred to as raw device mappings, or RDMs) to create a consolidated backup solution.

And before you ask, yes, there is a large difference between NetApp Snapshot technology and implementations based on copy-out snapshot technologies. In a NetApp environment, Snapshot copies do not move data and, as a result, have no performance impact on production storage.

In the early 1990s, Network Appliance revolutionized storage networking with a simple architecture that relied on NVRAM, integrated RAID, consistency points, and a unique file system to do things that the file servers of the time could not.

This technology is still the basis of every product that NetApp offers and includes:

- The WAFL® file system
- Snapshot copies
- Consistency points and NVRAM
- FlexVol™ and FlexClone™ technology
- RAID and RAID-DP

If you only read one paper about NetApp technology, read A Storage Networking Appliance.

Not All Snapshots Are Created Equal

Mercer Management Consulting found that NetApp environments require significantly less disk space than HP EVA, EMC CLARiiON, or EMC Symmetrix for the same size database.

The cause: different approaches to snapshot equivalent functionality.

Get the details. Read the full report. (See page 6 for a comparison of overhead capacity and disk capacity requirements.)

If you'd rather not take my word for it, check out the recent Mercer report [Decision-Maker Perspectives on EMC®, HP® and NetApp Storage Solutions in Enterprise Database Environments \(pdf\)](#). Page 6 documents how a typical database environment utilizing NetApp storage and Snapshot technology would require 1.44x overhead capacity versus 2.33x required with equivalent solutions from EMC and HP.

Q. Now you tell me. I've already invested in another type of environment and don't have the time or budget for a forklift change. Any suggestions?

Vaughn: Absolutely! NetApp provides a solution to backing up both physical and virtual machines with its Open Systems SnapVault® (OSSV) software.

OSSV is a unique solution in the backup industry as it provides individual system backups, which are stored on NetApp storage devices. OSSV provides a traditional file-level backup, yet it does so utilizing block-level incremental updates after the initial full backup has been completed. It can be used in either a local storage environment or with your existing SAN storage, and it enables you to significantly reduce the amount of data being backed up nightly.

You can combine OSSV software with storing your VMware environment on NetApp primary storage for a solution that provides instantaneous virtual machine restores and classic file-level recovery. Implementing OSSV results in a significant reduction of data transferred in and out of the ESX Server, thus allowing backups to complete quickly and eliminating the need for any artificial capacity limits. Restoration of complete VMs or individual files can be completed easily and in a traditional manner. In addition, OSSV is able to back up VMs that are not stored on NetApp storage systems, thus allowing for an easy integration into any existing ESX environment.

Then you can take that expensive third-party storage that you were using for backups and repurpose it for more demanding database or I/O-intensive applications.

Q. So how do I learn more?

Vaughn: I thought you'd never ask. You can find out more about all of these topics by reading [NetApp Storage and VMware ESX Server 2.5.x: Building a Virtual Infrastructure from Server to Storage \(pdf\)](#). This technical paper provides multiple examples and scenarios plus in-depth discussions on:

- Realizing thin provisioning and extending beyond physical storage limits
- Backing up virtual machines as physical servers
- Backing up virtual machines as files
- Enhanced backups with raw device mappings (RDMS)

Comment on this article



Tom Holland, senior engineer, Platform Hardware Engineering.
Tom has been a member of the NetApp Engineering team for six years and helped develop the F800, R200, and FAS6000 platforms. (left)

Naresh Patel, senior engineer, Performance Engineering.
Naresh has a Ph.D. in performance and has worked on platform and Data ONTAP® performance for seven years. (middle)

Howard Gage, manager, Hardware Engineering.
Howard is also a six-year NetApp veteran and was responsible for delivering the F800, FAS900, and FAS6000 series platforms. (right)

What Makes the FAS6070 so Fast?

An Engineering Roundtable on FAS6070 Storage Controller Performance.

The new FAS6000 storage family features two models, the FAS6030 and FAS6070. While the FAS6030 provides significant performance improvements, the FAS6070 has actually measured twice the performance of the previous high-end NetApp storage system, the FAS980 model.

Here's a quick comparison of how the two systems stack up:

	FAS6070 Active/Active Configuration	FAS980 Active/Active Configuration
Max. Raw Capacity	504TB	100TB
Max. Number of Disk Drives	1,008	672
ECC Memory	64GB	16GB
Nonvolatile Memory	4GB	1GB
SPEC SFS97_R1.v3 (operations/second @ ORT ms)	136,048 (w/ 1.09)	68,139 (w/ 0.98)

(source: [FAS6000 series tech specs](#), [FAS900 series tech specs](#), and the [SPEC SFS97_R1 results](#))

It's rare that a company releases a new high-end platform that provides a 100% performance improvement. To find out how NetApp did it, NetApp product engineer and best practice author Chris Lueth tapped some of the key engineers who worked on the new platform.

Chris: What types of goals did NetApp set for this new platform?

Howard: There were a number of dimensions. Of course we looked at raw performance, but we also focused on improving database workload performance with 64-bit addressability to support memory size greater than 4GB. NetApp has a unique advantage in its ability to provide true unified storage (that is, consolidating multiple types of storage and protocols on a single box: primary and nearline, CIFS and NFS, FC and iSCSI target mode, etc.), so we wanted to design in enough flexible I/O to allow customers to take advantage of that. We also decided to move to an embedded I/O to reduce cost.

Naresh: The other key area of focus was high-performance computing, which



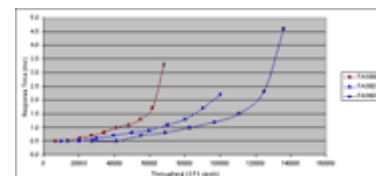
RELATED INFORMATION

- [Data Center Proven Launch](#)
- [FAS6000 Series Demo](#)
- [FAS6000 Series Tech Specs](#)
- [Architecting Storage for Resiliency](#)

SPEC SFS97_R1 Results

The Standard Performance Evaluation Corporation (SPEC) is a nonprofit corporation formed to establish, maintain, and endorse a standardized set of relevant benchmarks.

NetApp tested the new FAS6030 and FAS6070 systems using SFS97_R1.v3, the current SFS (System File Server) benchmark. The new FAS6070 system registered a 100% performance increase relative to the FAS980.



About NetApp SPEC results:

NetApp designs its benchmark performance test results for repeatability in real-world environments. Tests are set up with one volume to simulate a single entity and with RAID-DP™ turned on for protection against dual disk failures.

See full [SPEC SFS97_R1 results](#).

stretches the limits of most systems. In the past we had a relatively small number of clients simultaneously accessing data on our systems. Now we see chip vendors and computer animation studios that have thousands of Linux® servers and they often want to concurrently read/write large amounts of data from/to a single system. This means that the performance requirements of this system are growing considerably, so our goal is to improve overall price/performance by delivering excellent performance with the smallest number of systems.

Chris: Were there any particular technology considerations that impacted platform design?

Howard: Processor frequencies are becoming fairly static, so the trend is in adding processing power through additional processors and—as we go forward—through multicore processors. The challenge that we faced involved finding the right processor architecture to take advantage of cost-effective symmetric multiprocessing. We needed an architecture that offered 64-bit addressing, scalable memory bandwidth, and scalable I/O architecture.

Chris: What are some key architectural improvements to the new FAS6000 platform that allow the high-end FAS6070 to perform twice as fast compared to previous high-end NetApp systems?

Howard: Certainly the transition from older front side bus technology to HyperTransport™ links has provided tremendous internal and I/O bandwidth for the FAS6000 series relative to previous NetApp storage. There are four 4GT (Giga-transfers/second) links that operate concurrently in each FAS6070 controller, and at those speeds each link has tremendous burst bandwidth. If you compare the 4GT speed of the FAS6070 HyperTransport links to the 0.4GT on the FAS980, you'll get a feel for where some of the performance improvement comes from. That's an order of magnitude difference.

Naresh: In fact, the four HyperTransport links form a square interconnect with CPUs directly connected in each corner and deliver an aggregate burst bandwidth of 32GB/s per FAS6070 controller, or 64GB/s for an active/active, dual controller configuration. This is a huge improvement over the FAS980, which peaked at only 3.2GB/s and was limited by a shared bus. Having ten times the internal bus bandwidth of a FAS980 gave us a lot of headroom as we scaled up other components inside the FAS6070. This ranged from increasing NVRAM size to increased onboard I/O to the adoption of PCI Express (PCIe) technology for expansion cards.

One of the other things we wanted to address with this new platform is memory subsystem performance. You don't want multi-GHz CPUs spending a lot of time waiting for memory accesses to complete. A memory bottleneck that affects storage controller performance ultimately affects application performance and limits options for infrastructure improvements such as storage consolidation.

To prevent this, memory controllers are integrated in the CPU so that memory is directly connected rather than being routed through a slower intermediate chip. This reduces the time it takes to retrieve data from memory, ultimately using CPU cycles more efficiently. With industry-leading memory bandwidth and low memory latency, this is fastest hardware platform NetApp has ever designed.

As a result, the FAS6070 can maintain consistently high performance for applications ranging from intense transactional database workloads to very large e-mail deployments that cause pressure on memory. It facilitates consolidation projects by enabling different applications and workloads to be served from a single storage controller without starving the memory resources available to each. And in either case, with more memory and faster memory access, necessary background activities such as data replication and backups have considerably less impact on the performance of customer applications.

Tom: I'd also point out that the performance gains aren't only because of the increased internal bandwidth. With the FAS6000 series, NetApp has transitioned to a native 64-bit architecture in a platform that can take full advantage of it. In addition to operating on larger chunks of data in a single operation, the 64-bit architecture



Best Practices for Data Center Design

NetApp founder Dave Hitz and other experts share their perspectives on achieving:

- Maximum flexibility
- Consistent application service levels
- Massive scalability
- Highest application availability
- Highest staff/asset productivity



Watch the [TechTalk Webcast](#). Sign up for the [Proven Thursdays Webcast Series](#).

Architecting Storage for Resiliency

RAID-DP significantly increases data protection, with zero to minimal impact on capacity utilization and performance versus RAID4. And, because RAID-DP is an integral part of Data ONTAP, there are no hidden costs.

RAID-DP offers:

- Protection against up to two disk failures in the same RAID group
- Protection against single disk failure plus uncorrectable bit error during the reconstruction time frame
- No significant read, write, or CPU consumption differences
- Larger allowable RAID groups, which mean that capacity utilization stays about the same (one in eight disks dedicated to parity)

Read [The Private Lives of Disk Drives: How NetApp Protects Against Five Dirty Secrets](#).

Information-Optimized Backups

It used to be that a larger library, faster tape drive, upgraded network, or new backup server kept backups running.

provides a less obvious benefit to a high-performance system. With the 32-bit architecture of the FAS980, accessing more than 4GB of memory requires multiple operations; but on the FAS6070, all 32GB in each controller can be directly accessed in a single operation.

A good way to summarize would be to say that the FAS6000 is by far the best data flow engine we've ever designed.

Chris: Okay, lets dive into a couple of items a little further. How much more NVRAM memory does the FAS6070 have, and how does that help customers?

Tom: Whereas the FAS980 ships with 512MB of onboard memory on its NVRAM card, the FAS6070 has 2GB of memory on its NVRAM card. Since NVRAM is exercised more for write-intensive workloads, the extra memory on the FAS6070 NVRAM card helps customers see better performance for this type of workload. We have also changed the system interface on NVRAM to PCIe (PCI Express). This eliminates potential bottlenecks that the older PCI-Xbased slots might introduce. The new NVRAM provides adequate bandwidth to take advantage of the larger NVRAM memory during write-intensive workloads.

Naresh: In general, the larger NVRAM allows WAFL® to write to disk more efficiently, so customers see higher sustained write throughput from the system. Also, when there is a burst of writes in what typically isn't a write-intensive workload, say like a database, the extra memory on the NVRAM card helps absorb a longer burst of writes because writing to NVRAM is much faster than writing to disks. With the database example, if a database has logging activity at certain periods, then the resulting writes would go very fast, thus freeing the database engine for serving business activities.

In short, we increased the amount of NVRAM memory based on the high throughput capabilities of the FAS6070 to ensure optimal performance during writes.

Chris: That's a good segue into another one of the items we wanted to go into a little more detail on. What exactly is PCIe, and what is the customer benefit?

Howard: PCIe was designed to overcome bandwidth limitation issues with earlier PCI and PCI-X expansion slots. While the FAS6070 already has significant onboard I/O even without expansion cards, most enterprise deployments will likely populate many of the available expansion slots to accommodate their I/O needs. Since the FAS6070 was designed for high performance, we selected PCIe to take advantage of the bandwidth improvements it provides.

Chris: So how does the extra PCIe bandwidth translate into a tangible customer benefit?

Naresh: 100 MHz PCI-X slots are 0.8GB/s peak, and x8 PCIe slots are 4GB/s. PCI-X slots at 100 MHz could be shared between two slots, so a couple of fast HBAs could become limited by the PCI-X bandwidth. PCIe, on the other hand, is a point-to-point link, so each HBA has its own high-speed private path to memory. For customers, this means that they can add high-speed HBAs such as 10GbE and not have to worry about slot location; any available PCIe slot in the FAS6070 has plenty of bandwidth headroom.

Tom: In addition to increased bandwidth, PCIe provides improved RAS features. For example, instead of a shared bus, each link is point to point. A failed device on one HBA won't prevent the use of an HBA in another slot. Also, the protocol has improved error detection and includes the ability to correct errors if they occur on the link, so the system can survive certain errors that would be fatal to PCI-X.

Chris: Are there other things about the FAS6070 that you're particularly proud of?

Naresh: We did introduce a new 10GbE TCP/IP Offload Engine Ethernet adapter with the FAS6000 launch and as part of the Data ONTAP 7.2 release. The main reason for making this card available as quickly as we did was based on customer feedback about their needs to have higher throughput Ethernet infrastructure as well as to reduce their Ethernet port density. Many customers are transitioning from current 1GbE infrastructures to 10GbE, much like the transition from 10/100Mb Ethernet to 1GbE in years past, and they told us that they wanted 10GbE support in our storage products.

Along with increasing the Ethernet throughput with this adapter, we also provided

Over the past five years, however, disk capacities have grown 20x with a scant fourfold performance improvement in backup.

Today, many companies are turning to smarter backup solutions that:

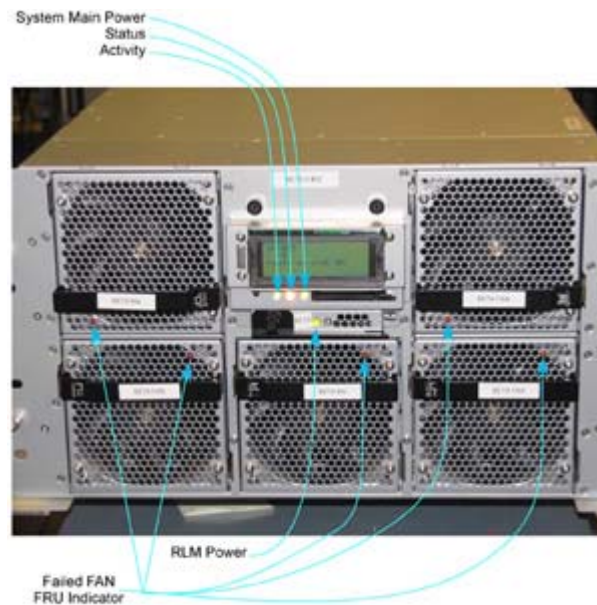
- Reduce backup storage consumption by eliminating redundancy
- Shrink backup windows by copying only incremental data changes
- Ensure reliable, recoverable backups by validating backups as they are created

Read [When Brute Force Isn't Enough: The Value of Information for Backups](#).

support for a TCP/IP Offload Engine capable card, also known as TOE, for this card and the 4-port 1GbE card. The TOE functionality basically saves CPU cycles on the storage controller by offloading most of the networking activity to its own processing engine, hence the name. Depending on the workload and storage platform, the TOE aspect of the card could provide significant performance improvement.

Tom: I'd definitely call out the failed FRU indicators. It's a fairly simple idea to provide LED failure indicators on the motherboard to highlight faulty FRUs. What's clever about them is that these LEDs stay lit up when the power is disconnected and the motherboard tray has been opened up.

I'm not aware of any other system where you can power the system off and an LED will keep blinking to identify the failed component. The goal behind this approach with the failure indicators is to help make serviceability much easier and faster.



Howard: Something else we should highlight is another new NetApp technology. The Remote LAN Module (RLM) is bundled with the FAS6000 series. The RLM is a standalone service processor inside the FAS6000 system that offers remote platform management of FAS6000. The RLM provides remote access to the system console over its Ethernet interface, using the secure SSH protocol. The RLM allows you to do remote troubleshooting, including power cycles and obtaining console logs, that just isn't possible with the serial console port. The RLM is also integrated with NetApp AutoSupport to call home and generate a customer case for any problems detected. We added the RLM to improve serviceability and manageability. (Learn more about [RLM](#) on the NOW™ customer site).

Naresh: One last thing I'd add is that the FAS6070 was designed for deployments where considerable scalability is required. The FAS6070 can handle up to 1008 disk drives and, when using the 500GB SATA disks, up to 504TB of storage. This type of capacity scaling enables the FAS6070 to support large numbers of clients, evolving customer application requirements, and multiple applications and types of storage on a single system.

Comment on this article



Darrell Mozingo

Systems Administrator, Synergy Data Systems

Synergy Data Systems is an IT outsourcing company providing a full range of services including software development, hosting, and server management. Darrell has spent the last six years working in IT. During his three years at Synergy, Darrell has been one of two people responsible for monitoring and managing all hardware infrastructure, including storage, servers, networks, and workstations.



Niraj Kumar Verma

Systems Specialist, Philips Semiconductors

Niraj has over nine years of IT experience and has spent the last six years working for Philips Semiconductors in Bangalore, India. Niraj manages the company's high-performance cluster and its UNIX® and Linux® servers and network environments.

Tool of the Month: Cacti

Every month, Tech OnTap showcases a free tool that just might make your life a little easier. Recommend a tool and get a FREE NetApp jacket. This month, two members independently nominated Cacti, so we've featured both of their perspectives.

Author: Cacti is an Open Source tool developed on SourceForge.net. The original developer was Ian Berry, with continuing contributions from many in the Open Source community.

What it is: Cacti is a complete Web-based front end to RRDtool, an industry-standard data logging and graphing application that can be used to write custom monitoring scripts and whole applications. Cacti includes advanced graph templating, multiple data-acquisition methods, user management features, and the ability to rapidly poll hundreds of devices.

How it works: Cacti stores the necessary information to create graphs based on user-defined specifications and populates those graphs with data from a MySQL database. Cacti is completely PHP Hypertext Preprocessor driven. Because Cacti has multiple data acquisition methods, it can monitor almost any networked device and provide an overview of network/server loads over any period of time. Cacti uses SNMP for gathering data, and it also allows the creation of custom scripts to gather data that is not available via SNMP. A large variety of scripts are available for devices of all kinds.

Why Darrell thinks it's cool: Cacti allows me to keep track of volume and inode usage on our NetApp storage systems, making addition and expansion planning a breeze. I can view usage and performance data over the past several years, months, or days to get an accurate and graphical view of what our network and storage systems are doing and what issues we may run into. The tool's templates make it easy to standardize monitoring and to quantify numbers for management.

How Darrell uses Cacti: We used to use MRTG for data collection and graphing, but it became too cumbersome given the number of devices we were monitoring. We now use Cacti to monitor everything on our network. In addition to monitoring two NetApp storage systems (an F87 and a FAS270), we use Cacti to continuously gather and graph data from over 100 servers (mostly Windows® servers with about a dozen Linux machines), plus all our networking equipment, including switches, routers, firewalls, load balancers, etc.

RELATED INFORMATION

Previously Featured Tools:

- [MRTG for Filers](#)
- [Nagios](#)
- [Simulate ONTAP](#)
- [SIO](#)
- [ToasterView](#)
- [NOW Tool Chest](#)
(password required)

Cacti Sample Screens:

The following screenshots show Cacti screens, including the dual pane tree view, device management, and preview mode.



Cacti allows us to standardize what we monitor using template groups that we've defined to include CPU, memory, disk usage, and other standard things. By applying a template group to a new server, we can automatically monitor all of these things. We also make use of a lot of data collection scripts with Cacti to gather specific information. Cacti connects to the specified device and polls information from the scripts. This is great because even though most devices support SNMP these days, SNMP doesn't always expose the data we want.

The information we get from Cacti helps us monitor trends, plan for upgrades, and quickly identify potential bottlenecks when troubleshooting problems. The graphical views of data make it simple to make the case to management for upgrades, additions, or other changes. If we need a different view of the data, we just generate custom graphs on the fly.

Cacti user forums such as forums.cacti.net and www.cactiusers.org are a great source for scripts to support almost any device you might have and to get answers to usage questions.

In addition to using Cacti for performance monitoring, we also use [Nagios](http://nagios.org) for ongoing health monitoring and alerts.

Why Niraj thinks it's cool: Cacti is very, very easy to implement and gives good and reliable results without affecting the performance of the devices being monitored or the server on which the application is run.

How Niraj uses Cacti: I've been using Cacti for several years to monitor things such as quota levels, volume space, CPU utilization, etc. for several NetApp storage systems in Bangalore, India. After downloading the NetApp SNMP Management Information Base (MIB) from the NOW (NetApp on the Web™) site, it was easy to configure the tool to provide trending data to create my growth plan. The tool gives me all the information I need when talking to management.

In addition to our NetApp environment, I use Cacti to monitor about 100 servers (200 CPUs) running a variety of operating systems, including Linux, HP, Solaris®, and Windows®. In our compute-intensive chip development environment, being able to quickly monitor all servers helps us to understand and manage long-running jobs.

In our company, different departments have their own servers. We use Cacti's access control feature to allow individual departments to directly view the performance data on their servers so they can check up on things themselves and we don't have to report performance data to them.

We also use 'thold' in conjunction with Cacti, which allows us to specify thresholds on various parameters such as CPU and disk utilization and to receive e-mail alerts when thresholds are exceeded. I downloaded this plug-in from the Cacti user [forum](http://forums.cacti.net). Incidentally, the Cacti user forums are very active, so you can get quick answers to any questions you might have as well as patches, plug-ins, and additions from the user community.

Caveats: According to Darrell, Cacti runs best in a Linux/Unix environment. The default polling program starts to choke up on 50 or so hosts; if you're monitoring more, you should replace it with the Cacti daemon (cactid). This program replaces the PHP-based poller in the default installation with a faster program written in C.

Download Cacti

Editor's Note: Tech OnTap previously featured [MRTG for Filers](#), a monitoring and graphing tool that has been modified to automatically query a NetApp storage system to discover its volumes and interfaces and to construct the configuration. MRTG for Filers is suited to situations where you need to monitor only NetApp systems. With its Web-based configuration and broader functionality, Cacti is probably preferable if you want to regularly graph data from dozens or hundreds of servers and systems where NetApp storage is just one component of the environment.

Want to recommend a tool? If we feature it, you get a FREE NetApp jacket.

 **Nominate a tool**

