

HIGHLIGHTS

- » **Why Protocol Should be Irrelevant**
- » **Take Back Your Disk Space**
- » **Improve Data Center Power Efficiency**
- » **Resources:**
 - **VMware® VDI Best Practices**
 - **How Oracle Does Consolidation**

The Future of NetApp: Riding the Technology Curves

Brian Pawlowski, chief technology officer, NetApp
Dave Ford, director of Advanced Technology, NetApp

In this exclusive Q&A, NetApp visionaries discuss the possible applications of new technologies, including hypervisors, pNFS, RDMA, and more, with a panel of Tech OnTap readers.

[More »](#)

BLOGGING WITH DAVE

Dave Hitz, NetApp Founder and EVP

"A LUN and a file are both just containers that hold blocks of data, so get over it. It was this similarity that originally convinced me that NetApp could 'unify' SAN, NAS, and iSCSI into a single storage appliance. (This is the key idea behind unified storage.)"

» [Dave's Blog](#)

DRILL DOWN

- » **NetApp Data Center Slideshow**
Take a tour of our Sunnyvale data center. See power efficiency from the inside out.
- » **Are You an Innovator?**
Be recognized with a NetApp Innovation Award for your innovative leadership or community contributions.



ADMIN RESOURCES

- » **VMware VDI Best Practices**
Get step-by-step operational guidelines for FAS systems with VMware Virtual Desktop Infrastructure.
- » **Optimize Your Oracle® Environments**
Hear the Webcast on how Oracle saved time and money by consolidating its database environment.

TIPS FROM THE TRENCHES

A Winning Strategy for Data Center Power Efficiency

Dan Hoffman, senior director of Facilities, NetApp

How energy efficient is your data center? Dan and Mark share techniques that help us exceed 2011 EPA targets today. Learn how we improve power efficiency by:

- Continuously eliminating inefficiency
- Challenging old assumptions
- Taking a team approach
- Attacking the problem on all sides
- Constantly monitoring and tuning all systems



[More »](#)

Oracle on NetApp Part III: Why Protocol Should Be Irrelevant

Steve Hight, director of Strategic Technology Projects, Catholic Healthcare West

In the final installment of this special feature, we bring you a case study of CHW running Oracle on NFS, FC, and iSCSI. CHW believes that unified storage—not a particular storage protocol—is the key thing.

[More »](#)

ENGINEERING TALK

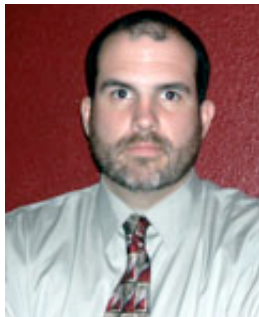
Maximize Storage Use with Space Reclamation

Richard Jooss, Senior manager, NetApp

Your Windows® servers may be using less disk space than your storage systems think they are. See how a unique NetApp feature helps you reclaim valuable disk space.

[More »](#)

[FEEDBACK](#)



Steve Hight

Director of Strategic Technology Projects, CHW

Steve Hight has spent over 15 years in the IT industry, the last 10 specifically in the healthcare field. During his seven years at CHW he has worked in some capacity on virtually every infrastructure-related team, including UNIX, Windows®, and storage—a professed passion. In addition, Steve also gained experience as a DBA and Java™ programmer, further broadening his technical background. As director of Strategic Technology, Steve puts this wide experience to use overseeing a variety of IT-related projects, including system and storage infrastructure, applications, and data center design.

Oracle on NetApp Part III: Why Protocol Should Be Irrelevant

By Steve Hight

The final installment of this three-part series on Oracle® storage architectures looks at a real-world deployment of Oracle on NetApp. Previous articles focused specifically on how NetApp engineering has optimized the use of [Oracle on NFS](#) and [Oracle on FC SAN](#).

When Tech OnTap contacted me about writing an Oracle on NetApp article, they asked me to talk specifically about our experiences running Oracle on SAN and NAS. I told them, "Shame on you for still thinking in terms of SAN or NAS. You should be talking about unified storage and using whatever protocol best meets each client's needs or the needs of the application." The terms SAN and NAS are dead along with "totally tubular" and "awesome." (Those are dead, aren't they?)

At Catholic Healthcare West (CHW) we have more than 117 Oracle Database instances running approximately 150 applications supported by 55 NetApp storage systems with NFS, Fibre Channel, and iSCSI protocols. In total, we have over 750TB of storage. Since the reality is that many IT organizations, application vendors, and DBAs still think in terms of protocol, this article highlights:

- How we choose which storage protocol to use for a given application
- The advantages of unified storage for Oracle
- The relative merits of the various protocols

Frankly, it's the second section that we think is the most interesting. The combination of unified storage and unique NetApp data management tools gives us tremendous power to manage our Oracle environment.

Choosing a Storage Protocol

Our first experience with Oracle on NetApp was with NFS a number of years ago. When NetApp said we could run Oracle over NFS, our first reaction was to laugh because at that time it was a very unconventional way to do things (some think it still is); back then our DBAs still wanted raw disk. When we converted from SAP to Lawson Financials we switched from traditional storage to NFS. Lawson sent a team

**OnDemand
TechTalk Webcast:**

**NetApp Solutions for
Oracle: Improve Data
Protection and
Response Times
through Consolidation**

RELATED INFORMATION

- [Oracle on NetApp Part I: NFS](#)
- [Oracle on NetApp Part II: FC SAN](#)
- [Best Practices for Oracle](#) (PDF)
- [NetApp Unified Storage](#)
- [Oracle Database Solutions](#)

About CHW

Catholic Healthcare West (CHW) is a network of 42 hospitals and medical centers in California, Arizona, and Nevada. CHW is the eighth largest hospital system in the nation and the largest not-for-profit hospital provider in California, with 8,500 doctors and a total of almost 50,000 employees and gross revenues of \$7 billion.

Choosing a Storage Protocol for Oracle

Your choice of protocol depends on a wide variety of factors that include existing infrastructure, processes, and skill sets, in addition to the relative capabilities of each technology.

NetApp offers a range of implementation options and tools for deploying and managing Oracle environments using NFS, Fibre Channel, or iSCSI, alone or in combination. Find out more specifics by reading the previous articles in this series:

of engineers out to certify the configuration. They planned to stay for five days. After one day they were completely satisfied; they certified the solution and left.

Since that time, we've added many additional Oracle applications, some on NFS, some on Fibre Channel, and some on iSCSI.

At this point you're probably wondering why we use more than one storage protocol with Oracle and how we choose which protocol to use.

Today, our standard for Oracle is Linux® with NetApp storage, and we don't differentiate between Fibre Channel, iSCSI, or NFS. We're running it all with confidence. When we deploy a new application, the decision really comes down to which protocol the vendor supplying the application can support the best. We always try to find out what protocol the vendor is most comfortable with or has the most deployments on. Some vendors are Fibre Channel only, while others are willing, or even prefer, to support other options. For example, our Emageon Picture Archiving and Communication System (PACS) specifies the use of NFS.

In the past, our preferred option was either iSCSI or NFS due to their protocol resiliency and the high per-port cost and management costs of Fibre Channel, but today the per-port cost on a director-class switch is really no longer an issue (it's no longer significantly different between FC and IP), so vendor support has become our overriding criterion.

Other Factors to Consider

Along with vendor support, there are a few other factors you may want to consider if you're trying to decide which protocol to choose for your Oracle environment:

- **Relationship with networking team:** I've talked with people who don't want to do anything but Fibre Channel because they don't have control over IP networking or a good relationship with their networking team. If that applies to you, you might find it beneficial in the long run to use FC SAN. On the flip side, there's obviously an issue of management complexity with FC SAN. If your organization doesn't have the skill set to deal with that, the protocols that leverage standard Ethernet gear may be preferable.
- **What does your DBA want?** Let's face it, good Oracle DBAs are hard to find, and many of them come with strong opinions about what's good and what isn't. Some will be more comfortable with NFS and iSCSI than others. I don't force a good DBA to use a protocol he or she isn't comfortable with if I don't have to.

Many people continue to believe that performance is a factor, but frankly I haven't found performance to be an issue with any of the protocols. Redundancy is excellent whichever you choose. We found that iSCSI has 2% to 4% more overhead versus FC. This probably resulted mostly from the fact that we were using a software initiator rather than an iSCSI HBA and in any case wouldn't be enough of a difference to sway our decision on protocol versus the other things I've already mentioned.

For us, the advantage of being able to use any storage protocol we need outweighs the relative advantages or disadvantages of any particular protocol. Read on to learn more about how we use unified storage in our Oracle environment.

Unified Storage in Our Oracle Environment

Over the years, CHW has grown by acquisition. As a result, we've got an extremely diverse environment with every flavor of UNIX® and countless applications. As a nonprofit, we have to be very standards based to mitigate cost and complexity; we focus on adopting technologies and deploying solutions that are open source, based on open standards, or at a minimum highly modular. It's just not a good business practice for us to adopt technologies that will leave us locked into a proprietary, single-vendor solution. As an organization, we are always looking for innovative and cutting-edge technologies to help improve patient care while reducing expenditures.

That's why unified storage fits so well in our environment. We can deploy a storage system and be certain that we won't need something different to meet our future protocol requirements. NetApp unified storage solutions meet our Oracle needs and give us a feature set that works across all protocols to simplify our Oracle data management.

- [Oracle on NFS](#)
- [Oracle on Fibre Channel](#)

IT Innovation in Healthcare

When you think about IT leaders, you may not immediately think of the healthcare industry. It may surprise you to learn that some of the most innovative IT teams that NetApp works with are in healthcare.

Faced with unique regulatory restrictions and tight budgets, these teams rely on advanced server and storage technologies to improve efficiency and ensure data availability around the clock.

Recent [NetApp Innovation Award](#) winners from the healthcare field include:

- [Sentara Healthcare](#) (2006)
- [Southcentral Foundation](#) (2006)
- [UZ Leuven](#) (2004)
- [BC Cancer Agency](#) (2004)

(Click the links to read recent feature articles on these IT leaders.)

See all the winners for [2004](#), [2005](#), or [2006](#).

A Case Study: Architecting Storage for Test/Dev

Ensuring the effectiveness of your development and testing efforts is critical, but the simple act of provisioning dev/test environments can consume vast amounts of storage and become a development bottleneck.

The company profiled in this case study first moved to NFS to simplify sharing and data management. Ultimately, it settled on a solution using NetApp FlexClone, allowing it to create space-efficient clones of production data sets in almost zero time.

[Read the case study](#) to get the full details.

Oracle Requirements

Our Oracle environment consists of literally hundreds of applications running on everything from Oracle 8.0.5 up to Oracle11g™ and Oracle RAC. A lot of people talk about their Oracle applications as "mission critical," but when you're in healthcare some applications truly are a matter of life and death. Our number one requirement from our Oracle infrastructure is reliability, followed closely by supportability. For us, it's very important to be able to support our storage in house without having to depend on outside parties. We don't like the typical glass house, big SAN thing here because we don't want to be dependent on others to help us recover from an incident.

A Few of the applications in use at CHW.

Application	Protocol	Purpose
FC SAN	Lawson Financials	Hospital's financials
Emageon	NFS	PACS application (radiology images)

Multiple Applications per Storage System

Right now, CHW has more than 50 Oracle servers (including three RAC clusters) connected to NetApp storage. Our 55 NetApp storage systems have a total of 750TB of data covering Oracle and other critical storage needs for our facilities. These storage systems are simple to configure and manage; we use NetApp Operations Manager to keep tabs on everything since we have a significant number of systems.

The majority of our storage systems are serving multiple applications. With the exception of our Emageon PACS application (radiology imaging) and Exchange, we don't have any NetApp systems dedicated to a single application. With NetApp flexible volumes (FlexVol® volumes), we don't have to worry about hot disks or disk contention. Workload is distributed very efficiently.

We'll be utilizing FlexShare™ in the future to prioritize workloads on the storage systems that carry multiple loads. We are not using it yet simply because DBAs and others are not generally aware that this capability exists. Once word gets out, we should have a lot of good applications for the technology.

Storage System Reliability and Rapid Recovery

With Oracle on NetApp, we know the storage is reliable, and it's also very easy for us to recover if database corruption occurs without any outside assistance. We can always use SnapRestore® to revert back to a point in time before the corruption and get back up and running in a matter of minutes. In addition, we like the ease with which we can clone a database using FlexClone® for test/dev or other purposes without consuming twice the storage resources.

Thin Provisioning

We're also using NetApp FlexVol technology to enable thin provisioning. Almost every time a new Oracle application (or other project) comes online, we see extremely bloated requirements for how much storage a database is going to need. Using flexible volumes enables us to mix volumes with differing performance and capacity needs and grow and shrink them as requirements change. We can tell a customer, "Sure, I'll be happy to give you 1TB for your new project," and know that space isn't going to get consumed unless it's actually needed. Thin provisioning allows us to overprovision our storage and allow volumes to grow as necessary rather than allocating large volumes up front and having them sit mostly empty for the first year.

Data Protection

In my experience, most data protection software is hard to configure, carries a steep learning curve, and often just doesn't work as advertised; as a result it ends up as shelfware. That's why we rely so much on Snapshot™ and SnapRestore. It's very easy, people use it, and people trust it.

Our data protection environment utilizes a combination of NetApp Snapshot copies for immediate recovery needs and regular tape backups. Now that SnapManager® for Oracle is available we are considering using it for new deployments. But as we all know, it takes time to move from a scripted solution to a product solution, but we see the benefits of having a product to simplify this process and the support from a

vendor. We will get there eventually!

Observed Protocol Differences

Although we are agnostic on protocol, obviously there are some differences. As I said previously, these differences aren't enough to make us pick one protocol over another, but they may be worth considering if you are choosing a storage protocol for your environment.

We like NFS for its ease of management and statelessness. We can recover a 200GB database in a minute on NFS, which is impressive. On the other hand, fewer vendors support it, and NFS is less consistent across platforms. Different client platforms require different patch sets and different tuning for optimal performance. Vendor patches have been known to break NFS. It appears that direct NFS will address many of these issues. It's a technology we're actively investigating for future deployments.

Fibre Channel, on the other hand, is the mainstream protocol. It's more consistent across platforms, predictable, and robust. It does, however, add management complexity, and provisioning is somewhat slower, although with NetApp it doesn't make a huge difference.

The iSCSI protocol is very easy to manage and as resilient as NFS. We manually fail over our NetApp clusters on iSCSI to perform maintenance on occasion, and, just like NFS, we don't incur any downtime. The failover happens fast enough that the applications don't go down. I am continually surprised to meet people who are not familiar with iSCSI, so it would seem that iSCSI is not yet as mainstream. We have investigated using hardware-based iSCSI initiators, but decided against their use simply because the minimal CPU overhead does not justify the cost of the cards and the additional management overhead.

Some observed advantages and disadvantages by protocol.

	Advantages	Disadvantages
FC SAN	• Robust and predictable • Broad industry support	• Complex management • Provisioning is slower
NFS	• Easy to manage • Stateless • Exceptional recovery	• Patching and tuning* • Different on each platform • Fewer application vendors support it
iSCSI	• Free • Stateless • Low per-port cost	• Support not as broad as FC • 2% to 4% CPU overhead

*Patching and tuning are no longer necessary with the direct NFS client in Oracle11g.

Getting the Most Out of Storage

At CHW, we prefer to buy fewer, larger storage systems and use each system for multiple workloads. Unless you have extreme performance requirements, there is no longer any reason to dedicate a storage system to one application. Today's storage systems are more than fast enough to serve multiple applications in most cases. By choosing unified storage that supports NFS, CIFS, Fibre Channel, and iSCSI we can use whatever protocol makes the most sense for the application and change protocols whenever necessary. Unified storage gives us the flexibility to meet all our storage needs for Oracle and other critical applications.

TECH ONTAP ARCHIVE



Richard Jooss

Senior manager, Product and Partner Engineering, NetApp

Richard Jooss is senior manager of SAN Product and Partner Engineering at Network Appliance. Rick is responsible for defining technical and business requirements for the SAN ecosystem and storage, and business solution integration with NetApp SAN solutions. Rick has 15 years of experience in the storage industry. He holds a bachelor of science degree in Electrical Engineering and Computer Engineering from the University of Wisconsin.

Reduce Storage Costs by Reclaiming Disk Space

By Richard Jooss

Using the storage space you already own more efficiently can be a great way to get control of your storage environment and reduce your storage expenditures. NetApp has pioneered a variety of technologies designed to reduce data management overhead and limit the amount of storage you need:

- Space-efficient **Snapshot™** copies
- **FlexClone®** technology to clone storage volumes without using 2X the disk capacity
- **Thin provisioning** options that reduce the amount of storage you need to allocate up front
- **A-SIS deduplication** technology that eliminates duplicate copies of data blocks within any NetApp storage volume

This article examines a relatively new NetApp technology—**space reclamation**—that helps you conserve and re-use the disk space you've already allocated to LUNs in Windows® environments. NetApp is the only vendor to offer this capability.

Understanding Space Reclamation

To understand space reclamation, you have to remember that when you assign a LUN to a host, the host creates its own file system on the LUN and does its own bookkeeping. As a result, the host can have a very different idea of how much space it is currently using within the LUN than the storage system has. (This applies equally to all SAN storage vendors, not just NetApp.)

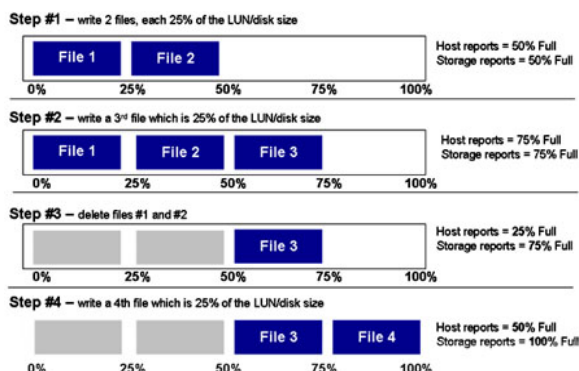


Figure 1) How host and storage views of space usage get out of sync in reclaiming space.

RELATED INFORMATION

- [Thin Provisioning in a NetApp SAN \(PDF\)](#)
- [Nick Triantos' Blog](#)
- [SnapDrive and SnapManager®](#)
- [A-SIS Deduplication](#)
- [Storage Efficiency \(PDF\)](#)

Thin Provisioning for NetApp SAN Environments

One of the big disadvantages of traditional disk arrays is that they force you to allocate dedicated storage space to a disk volume or LUN when you create them. Since it's often hard to gauge the amount of space you'll need up front, you end up overprovisioning and wasting valuable disk space.

In contrast, when you create a LUN on a NetApp system, you don't have to dedicate specific disk blocks. Instead, blocks are allocated only as data is written. In this way, multiple LUNs can flexibly share the same pool of free storage. You simply add more capacity when free storage gets low, and you can painlessly grow a LUN if more space is required.

Learn more. [Read the technical report.](#)

You'll want to refer to Figure 1 during the following explanation. Consider a new LUN that has been allocated to a host. In Step #1 we write two new files, each consuming 25% of the LUN. Both the host and storage report that 50% of the space has been used, as expected. If we add a third file of the same size (Step #2), again the host and storage report the same utilization.

But now suppose we delete the first two files as shown in Step #3. Remember that with most file systems (including NTFS), deleting a file causes the file system to deallocate the blocks and put them on a free block list. Nothing happens to notify the storage system that anything occurred. The data stored inside the LUN is opaque to the storage system. Hence, the views of the host and the storage system suddenly diverge. The host reports that the file system is only 25% full while the storage system believes that 75% of the LUN is in use.

The host is under no obligation to re-use the blocks it just freed, so if another large file is subsequently written to the LUN, it may occupy previously unused space, as shown in Step #4. The storage system now thinks the LUN is full, while the host shows just 50% utilization.

These discrepancies don't really pose a problem in situations in which you've dedicated a LUN of fixed size to an application. But there are two places where it matters:

- **Thin provisioning.** If you use thin provisioning, there's a big potential difference on the storage system between a LUN that is considered 25% full and one that is 75% full.
 - Suppose your LUN is 2TB in size. In this example, that would be equivalent to a full TB of storage you could use someplace else.
 - In the worst case, if you have a thinly provisioned volume configured to "autogrow" on your storage system and the storage system shows the volume as full, it could allocate additional space even though the host view of the LUN shows plenty of free space.
- **Snapshot copies.** Guess what? If the storage system doesn't know the blocks allocated to deleted files aren't actually in use, it's going to preserve them in Snapshot copies. Over time you may need more storage space for your Snapshot copies (a larger snapreserve, as it's called at NetApp).

Reclaiming the Unused Space in a LUN

At this point, you can see that there are advantages to reclaiming storage space. When space reclamation starts, an agent on the host determines which blocks on a LUN are not in use by NTFS. The host agent then communicates this information to the storage system and WAFL® marks the blocks as free. Once the blocks are freed by the storage system, they are no longer retained when Snapshot copies are made. If thin provisioning has been enabled, the space is once again available for use by other LUNs.

Without space reclamation, it is almost impossible to thin provision a LUN over a long period of time. Naturally, this depends on how the particular application uses storage space, but in general the storage system tends to allocate more and more of the space until the LUN reaches its maximum size, at which point it is the same as if you provisioned the LUN normally. Figure 2 illustrates how space reclamation recovers the space.

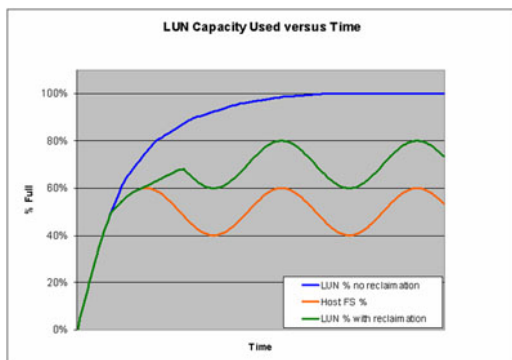


Figure 2) Effect of space reclamation on LUN space usage.

SnapDrive and SnapManager

SnapDrive offers a number of features besides space reclamation that greatly simplify data management tasks. In a recent Tech OnTap article, NetApp technical director Kostadis Roussos described the evolution and technical capabilities of SnapDrive as well as the suite of SnapManager tools for application integration.

[Read the article](#) to learn more about the inner workings of these valuable tools.

Deduplication Comes of Age

Technology that helps you recover disk space you've already allocated. A-SIS deduplication can help you get back huge amounts of disk space by eliminating duplication.

At its heart, A-SIS deduplication relies on the time-honored computer science technique of reference counting.

When A-SIS deduplication is enabled on a volume, it computes a database of fingerprints for all of the in-use blocks in the volume (a process known as "gathering"). Once this initial setup is finished, the volume is ready for deduplication.

To avoid slowing down ordinary file operations, the search for duplicates is done as a separate batch process. As the file system gets updated during normal use, WAFL creates a log describing the changes to its data blocks. This log grows until one of the following occurs:

- The administrator issues a sis start command
- The next time specified in the sis config schedule occurs
- The changes to the log exceed a predetermined threshold

Learn more. [Read the article.](#)

Space reclamation may be particularly useful in any environment that shows a large variance in the amount of data used. Systems used for queuing print jobs, e-mail, or standard file sharing are excellent candidates for space reclamation on a periodic schedule. Any system that has gone through a major change is also a good candidate for space reclamation. If you've performed database reorganization or made significant changes to a database because of archiving, space reclamation will likely recover a significant amount of data.

Getting Started with Space Reclamation

NetApp space reclamation technology has been integrated into NetApp SnapDrive® for Windows 5.0 and comes at no additional cost. SnapDrive makes it easy for Windows server administrators to manage NetApp SAN storage using either a wizard-based approach or through a command-line interface (SDCLI) that can be used in conjunction with management scripts and scheduling programs.

Using the SnapDrive GUI or SDCLI, you can periodically initiate the space reclamation process on your LUNs. The GUI tool will first determine how much space can be reclaimed and ask if you wish to continue. You can limit the amount of time the process will use to ensure that it does not run during peak periods.

Here are a few things to keep in mind when you run space reclamation:

- It's a good practice to run space reclamation before creating a Snapshot copy. Otherwise blocks that should be available for freeing will be locked in the Snapshot and not able to be freed.
- Because space reclamation initially consumes cycles on the host, it should be run during periods of low activity.
- Normal data traffic to the LUN can continue while the process runs. However, certain operations cannot be performed during the space reclamation process:
 - Creating or restoring a Snapshot copy stops space reclamation.
 - The LUN may not be deleted, disconnected, or expanded.
 - The mount point cannot be changed.
 - We don't recommend running Windows defragmentation.

Get More from Your Storage

With space reclamation, NetApp gives you a new tool to further increase storage efficiency. Space reclamation makes it possible to efficiently thin provision your block-based LUNs for long periods of time without having LUN space allocation slowly creep up to 100%. It also increases the efficiency of Snapshot copies on your LUNs, reducing the amount of space you need to reserve for this important function.

TECH ONTAP ARCHIVE



Brian Pawlowski (aka Beepy)

Chief Technology Officer, NetApp Engineering

Beepy (left) joined NetApp engineering in 1994 as employee 18. Brian's interests span a broad range of topics, and is especially passionate about virtualization, pNFS, and IP storage protocols.

Dave Ford

Director of Advanced Technology, NetApp Engineering

Dave was engaged in joint development work around storage networking using RDMA technologies such as virtual interface architecture (VIA) prior to joining NetApp via one of the company's first acquisitions in Orca Systems.

The Future of NetApp: Riding the Technology Curves

An Experts Tech OnTap Member Interview with NetApp Engineering

Recently, Tech OnTap hosted a roundtable discussion between members of the NetApp Advanced Technology group and the Tech OnTap Member Review Board. (See sidebar for review board profiles.) Topics covered include virtualization, hypervisor technology, pNFS, RDMA, Xen, search, disk reliability and performance, NetApp platform and OS convergence, and much more.

This interview describes Advanced Technology investigations at NetApp. There are many forward-looking statements, and the material is by its nature speculative. NetApp makes no commitment to any product deliverables based on the information provided here.

Specific topics addressed in answers to panel questions:

- [Advanced Technology group overview](#)
- [Technologies and trends](#)
- [Hypervisor technology](#)
- [pNFS](#)
- [RDMA](#)
- [Reconstruction time for large disks](#)
- [Convergence of FAS and NearStore® \(impact of Data ONTAP® GX technology\)](#)
- [Software-only solutions](#)
- [Future of data management](#)

Tech OnTap: Before we talk about specific technologies and projects, could you provide a quick overview of the NetApp Advanced Technology group?

Beepy: The [NetApp Advanced Technology group](#) was started more than two years ago as a separate group within Product Development. In addition to ongoing advanced development work within product development teams, NetApp saw a need for a group specifically tasked to handle longer term speculative projects, or projects that were very broad or cross-functional that spanned several Product Development

RELATED INFORMATION

- [Dave Hitz: Future of Storage](#)
- [The Core NetApp DNA](#)
- [NetApp Vision for Management: From Reactive to Proactive Admin](#)

Advanced Technology Group Interview: Customer Participants



**Tony, Senior Storage
Engineer at a Large
Software Company**

"I manage approximately 30 NetApp storage systems scattered around the world. We're using them mostly for file storage, and we're just moving into the Oracle® Database sphere. We are also using NetApp to host multiple VMware® instances."



**Reinoud, IT Manager,
Leuven University
Hospital in Belgium**

"I work for one of the largest healthcare providers in Europe. We've been using NetApp technology since 2000, and today we run all our storage on NetApp. We use it not only for our mission-critical Sybase environment but also for VMware, our PACS environment, and home directories. Today we have more than 270TB, and last week I ordered

groups.

For instance, when the RAID group was looking at new RAID algorithms, some of the advanced development work for RAID-DP™ was done within the group. But for things that span many groups, it makes sense to factor out and take a broader, more strategic approach with a focused group that reports in to the director of ATG, Dave Ford.

Dave: The Advanced Technology group is really one aspect of a portfolio of investment in advanced development work. As you all know, we have a very innovative product and still continue to add tremendous enhancements. What ATG is doing is looking out 18 months to as much as five years—beyond the next couple of product releases—to do speculative technology development and research that may be useful in the future.

Beepy: I think one of strength of NetApp has been its practical approach to delivering products to customers. Our approach is to come up with a blended model that balances our investment and places bets along several lines. David and I have been working to keep the advanced technology program grounded while still being speculative: working to take themes that come from customer problems and drive them forward and answer specific questions that could lead to potential products in the future.

Steve Kleiman, NetApp chief scientist, came up with a thinking tool several years ago for how senior technical leadership should drive innovation forward; he called it the "3 Cs":

- **Customers:** Particularly customer problems that remain to be solved, plus emerging customer application requirements.
- **Competition:** The reality is that we're out there competing with other companies' products, and there's sometimes a horse race going on in terms of who's doing what. This can trigger a response in our product set.
- **Technology curves:** These are the trends that drive change in the industry. For example, one trend that every storage vendor must deal with is the fact that disk drives are increasing rapidly in capacity while their performance remains relatively constant.

Tony: What are some of the things that are really exciting in terms of technologies that you're working on and what trends are you looking at right now?

Dave: One area is virtualization, specifically the benefits that hypervisor technology may provide us for not only building storage systems but also other products. We've invested in indexing and search technology with the goal of integrating such capabilities into the storage system. This would be targeted to enterprise storage search, versus your typical Google solution, which is targeted toward end users. This has been a very successful investment for us and is already moving into product development. [See the final question in this article for a more complete discussion of index and search, including the IS1200 and Kazeon. Ed.]

Another area where we continue to advance the bar is in NFS in multiple ways; pNFS [parallel NFS] was actually our first ATG project that we started back in September of 2004, when we hired Garth Goodson from Carnegie Mellon. Garth's work is helping to drive the pNFS specification (part of NFSv4.1) to completion. NetApp also recently contributed a sample Linux® implementation of pNFS.

Another project under way is looking at ways to deliver a heterogeneous federated namespace for NFS. This extends to the ability to cluster Data ONTAP GX nodes together in a heterogeneous fashion with solutions from other vendors such as IBM. We have been collaborating with IBM research on this project.

We've also been investigating remote direct memory access (RDMA) technology for increases in scalability, improved clustering for Data ONTAP GX, and other benefits that it may offer for storage networking, especially NFS and iSCSI. We've created a reference implementation of NFS/RDMA (both client and server) to accelerate

another 240TB, which also includes a four-node Data ONTAP GX cluster."



Rajeev, IT Manager at a Leading Informatics Life Sciences Institute

"Our primary research interests revolve around the advancement of the science of genomics. This type of research places an incredible demand on grid computing and storage infrastructure. Our grid computing environment currently occupies 630 slots and has over 1TB of memory. Our storage environment includes two FAS6070 clusters and about 100TB of usable storage. We were an early Data ONTAP GX beta site, and we are currently investigating related technologies such as deduplication and virtualization."

Advanced Technology at NetApp

The Advanced Technology Group (ATG) is one aspect of a portfolio of advanced technology (AT) investment at NetApp.

ATG is tasked with looking out 18 months to as much as five years—beyond the next several product releases—to do speculative technology development and research that may be useful in the future.

Support for AT projects within NetApp Product Development is the second aspect of the portfolio. The Advanced Technology group and its associated AT council—Dave Ford, Brian Pawlowski, and Chief Scientist Steve Kleiman—act as an internal venture capital group, soliciting proposals from the product development community and funding those that have business value as AT projects within Engineering.

The third aspect of AT investment is university research. NetApp funds research at leading universities that are doing research in storage and systems, including Wisconsin, Santa Cruz, Harvard, Carnegie Mellon, and others.

Looking at the complete portfolio, ATG is a riskier investment with a longer term return. Advanced technology work in Engineering gives us more near-term returns and not as much risk, while university research is the most speculative part of the portfolio.

adoption in the Linux kernel community.

We are also looking at ways to continue to scale system performance for block-based workloads. We think that there are some fundamental things that we can do to further enhance our position in the SAN space.

Beepy: Here's my thesis about why we need to play with virtualization, especially from the customer's perspective. I firmly believe that Network Appliance has the best storage solution today—and in the future—to support application server virtualization because of the way we view data and our fine-grained storage provisioning policies around FlexVol® volumes, Snapshot™ copies, cloning, combined with our ability to do these operations rapidly. So, my vision for application server coupling, in terms of the use of hypervisors and virtual machines on application servers, is getting more and more interesting features in our administration framework to support server virtualization. For instance, this could allow rapid provisioning of a 5,000-node Linux cluster starting from a single image using shared, read-only images that are consuming little or no space because you simply clone the image over and over and over again to rapidly fire up a compute cluster. NetApp [FlexClone®](#) makes this extreme level of space efficiency possible.

Our underlying storage virtualization technology supports emerging trends in server virtualization, and we're looking to exploit it, both from a research perspective—one of the relationships we have is with Cambridge University for the Xen open source virtual machine work they're doing—and from a product perspective.

Rajeev: Where do you see hypervisor technology converging into what you're trying to do now

Beepy: From my perspective using hypervisor technology underneath Data ONTAP is the most advanced work we are doing. About two years ago, I was sitting in a pub with Ian Pratt of Cambridge University, and I said, "So people are using this Xen stuff successfully?" He said, "Absolutely, and we know that other customers out there are using different virtualization solutions." I said, "Is there anybody shipping an enterprise storage system with a virtualization hypervisor underneath?" He replied, "Not yet."

Hypervisor technology is still emerging in the storage server market. We're trying to get ahead of the train and figure out how to exploit it as it goes mainstream. It provides opportunities for increasing the availability of a storage device by providing increased fault isolation. I took these ideas back to the AT team and said, "Show me how hypervisors can add value in terms of fault isolation, fast restartability, and minimizing application disruptions over and above what we already do in Data ONTAP today." So the hypervisor project is going down that path.

Dave: There are many benefits of hypervisor technology that we're evaluating, and I'm just going to add one more. One of the possibilities is to support third-party, value-add applications in a protected Data ONTAP domain. We think that this has potential and will allow us to more easily extend Data ONTAP capabilities without doing the software equivalent of brain surgery.

Rajeev: I was wondering if you can go into a little bit of detail regarding pNFS.

Beepy: I just came back from the Prague meeting for the IETF, and we're in the last stretch of NFS v4.1, the follow-on to the core NFS v4 protocol. pNFS is a big feature in 4.1, although this will take some time to get out in terms of product. But now we have a protocol specification for people to implement to.

The cool thing about pNFS is it provides the ability to do bandwidth aggregation on the front end, much like a clustered file system approach but in a standard protocol that will be as widespread as NFS. Instead of having to pick and choose among various noninteroperable cluster file system solutions, you'll be able to deploy pNFS instead. What's really cool about pNFS is that it meshes very nicely with FlexVol HPO—the striped FlexVol in Data ONTAP GX.

The pNFS client provider gives you the ability to stripe the data across multiple NFS servers. pNFS provides for bandwidth aggregation across the network from an

The Core NetApp DNA

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

These basic building blocks of NetApp technology support:

- Multiprotocol environments (NFS, CIFS, FC, iSCSI, and so on)
- Clustered failover, mirroring, and disk-to-disk backup
- RAID-DP and other software-based resiliency features
- The near-instantaneous creation of writable clones
- Block deduplication using A-SIS

Learn about [The Core NetApp DNA](#).

application server or cluster of application servers through the network leading into the storage architecture. FlexVol HPO is our data layout engine for pNFS. It provides bandwidth aggregation on the back end across a number of nodes. Together, the two provide you full bandwidth aggregation to disk. It's a match made in heaven.

Tony: Could you elaborate on RDMA and how it relates to storage networking?

Dave: Network Appliance has always been on the leading edge of advanced interconnect technology. We were the largest consumers of Gigabit Ethernet when it first appeared, and we were one of the first companies to start working with ServerNet technology, and we were the first commercial vendor to ship Infiniband. We're already large consumers of 10 Gigabit Ethernet. A future use will be to maintain distributed state between controller pairs which at some point will go to an N+1 configuration. The interconnect is in the critical path for I/O response time. So we have a very critical need for high-performance interconnect technology; it's really fundamental to what we do.

One of our objectives with RDMA technology is to invest in an industry-supported software and hardware solution that will allow us to exploit these fast interconnects, while being more flexible about the specific vendors or transports that we can work with. There's an RDMA stack that's being pushed by an open source community, the Open Fabrics Alliance, and this supports not only Infiniband but also RDMA over Ethernet, called iWarp. It's supported by multiple vendors. The idea is if this will work for us, then we don't have to invest in our own proprietary driver stack, and we don't have to work with vendors to have them craft a custom solution for us. We can better ride the commodity curve for high-performance interconnects.

Reinoud: As disks get larger, reconstructions take a lot longer. Even with RAID-DP, reconstruction may not be fast enough with disks of 1TB and larger. Are there other options for RAID protection than just mirroring every disk? What else is changing in terms of disk reliability and performance?

Dave: As we get more and more data per spindle, we need to do something different from a performance perspective. One way to solve that problem is, of course, to stripe the data across more and more drives, but that raises concerns about reliability. Is RAID-DP going to be sufficient? One thing we're currently investigating is ways get more drives working on reconstruction in parallel, which helps solve this problem.

A further method we're investigating is to be able to stripe data not only across more and more drives but also across multiple nodes. That's something that we're doing with Data ONTAP GX, and we'll continue to exploit this capability. Is there something that we need to do at the node level to insure better reliability, availability, and, of course, disaster recovery beyond simple mirroring of data? Can we do something more intelligent at a finer granularity at the node level? These are questions we continue to ask ourselves.

Beepy: There is a first principles approach to dealing with large disks. You don't want to be in an unprotected mode in a single RAID solution while you're reconstructing. So enter RAID-DP: protection against double disk failures. I think the important thing about the RAID-DP solution (in case no one's noticed): it's not that we did it; it's that it's on by default when you create new volumes on your storage systems today. We implemented it in a way that was efficient to the point where it can be the default. This was the key insight in the design and the implementation of RAID-DP.

Next you ask, "What can I do as an engineer to make this solution practical, and how do I make storage systems that work for customers, and how do I minimize the impact of any possible disruption in service?" Maintenance Center is a feature of our product set today. The Maintenance Center does real-time diagnostics of disk drives to keep drives running proactively by monitoring their performance during operation. We expect increased investment in clever ways to provide continued service and reliability in our storage systems by making our devices more self-diagnosing and self-healing. We recently hired a senior engineer with this specific background into the AT group.

Beyond things like RAID-DP, observations such as "disks often don't fail completely" allow you to avoid doing a complete parity reconstruction in many cases. You can copy all of the correct data and then factor out just the small portions of possibly incorrect data that resulted from a transient error and have a very efficient algorithm for drive reconstruction that doesn't involve doing parity reconstruction over the entire drive with all the drives in the RAID group. Our ultimate solution includes both this practical engineering aspect coupled to the theoretical aspects through technologies like the RAID-DP algorithm.

Tony: I'm curious about the convergence of the FAS and NearStore personalities. Can you talk about that?

Beepy: Where we are today with the two platforms emerged from a technology trend that Steve Kleiman drove very hard at NetApp, the emergence of very large, low-cost disks. These emerging dense drives presented an opportunity to provide large low-cost storage for certain applications that weren't necessarily in the primary data center. Our NearStore products, which we defined as a market that was separate from primary storage, emerged from this trend analysis.

I almost had one of those heart-stopping moments when I was on a customer visit in Spain with Dave Hitz—who certainly is allowed to say whatever he wants to say, given that he started NetApp. We had just introduced the NearStore product, and he was encouraging customers to start deploying NearStore in primary storage applications, which was in conflict with our product marketing message at the time. His point was that commodity technology that emerges in a particular application, such as near-line storage, eventually will go mainstream. That this is just the inevitable way that technology works; it finds a toehold and then makes its way into standard deployments.

We knew that eventually our primary data storage and near-line storage technology would become one. Now we have systems with ATA drives or ATA-like drives—various technologies are coming in the future that bridge the gap between ATA and Fibre Channel drives—and the lines have become blurred. Moving forward, I think we'll use the NearStore personality to drive new market segments outside what one would consider primary data storage.

When you look at a large database deployment, you typically want the highest performance drives. Basically the cost factor is not as important as the availability and the performance requirements. Then there's going to be another segment of your data that involves compliance that still requires reliable data handling and data integrity, but where you'd very much like to minimize the cost. That's where you can best exploit power savings and also where you have large amounts of accumulated data that you need to search and index. These are the areas where the NearStore personality fits.

There's also the convergence of the Data ONTAP GX technology with the rest of our technologies. The Data ONTAP GX technologies, with horizontally scalable clustering for transparent load balancing, plus migration and replication, allows us to provide the data management features that you know and love with Data ONTAP in a horizontally scalable solution that grows transparently with your application.

The interesting thing about Data ONTAP GX for this discussion on primary and secondary storage is that the Data ONTAP GX clustering technology is going to make the difference between primary and secondary storage transparent from the perspective of accessing the data and naming the data from the application. With its global namespace and transparent migration capabilities, Data ONTAP GX gives you the ability to move data—between storage types, for example—without disrupting applications.

Your application data goes through various phases—for instance, when you have a software product in development, your development team is working on it actively, and it's kept on primary storage. Once you release the software, there are maintenance activities, but the source is going to be accessed less and less frequently as it ages. So this data moves through various parts of your storage from primary storage to secondary storage and into archive mode.

Using Data ONTAP GX technologies and adding some search capabilities and some

policies regarding how to manage data, you could keep the data available from an application perspective, while giving the storage administrator a lot more flexibility about where the data resides to optimize the cost of deploying storage. The path always remains the same, while the data is moved transparently on the back end.

We are focused on enabling customers to optimize storage deployments in terms of cost as technologies converge, and we're working on ways to simplify data management through policy-based migration that reduces or eliminates the need for storage administrators to manually move data.

Rajeev: Could this make NetApp move away from the combined hardware/software solution you have today toward a software-only solution?

Beepy: When James Lau and Dave Hitz founded NetApp, it was their intention to have a software-only solution. They were going to have a number you call, and you just download some software and install it on whatever hardware you were running, and it would just work. What they found was that the reliability and availability of the system were not simply a software problem. They had to do work in hardware, they had to qualify drives, and they had to qualify motherboards to provide a solution that would deliver the necessary reliability. Customers were expecting a solution that included the storage itself.

We've asked ourselves again whether there's a viable storage solution other than a hardware/software solution like we do today. We're not there yet, but I find the question interesting because one of my fascinations with the hypervisor technology is the way it levels the playing field in terms of a common architecture. I think it provides a lot more ability to start flexing around a software solution compared to being tied to a specific hardware instance.

The thing that Data ONTAP GX has—besides horizontal scalability and transparent migration—is a software architecture that is highly modular. As David pointed out, we're looking at hypervisors for running third-party software applications on top of our product in a protected sandbox, so they won't bring down primary storage service while extending Data ONTAP capabilities. It may be that we can do certain pieces of a storage architecture as a software solution that runs in concert with part of our product being provided in a similar form as it is today. That might be the first step toward a pure software solution that you can layer onto a grid.

Rajeev: We're testing Kazeon right now. We may have solved the problem of deploying and provisioning storage, but I don't think the storage industry has solved the management problem. NetApp seems to be on the right path. In my mind I stitched together Kazeon plus SRM-type policies, plus Data ONTAP GX. Is that what you're thinking about?

Dave: Kazeon is a very good partner, and it has excellent technology around indexing and searching of file content. We did some initial work in this area as well and decided that Kazeon has a better approach, and it's already available in the marketplace.

One of the more fundamental investigations that we're doing is around metadata indexing. We see a scenario where you would use Kazeon for classification, while NetApp does metadata indexing for automated data management. By integrating this with the storage system, we think we can improve the performance over a third-party, appliance-only solution.

TECH ONTAP ARCHIVE



Dan Hoffman

Senior Director of Facilities, NetApp

Dan Hoffman has over 20 years of experience in facilities management, with expertise in cogeneration and energy conservation. He currently manages the data center operations team at NetApp headquarters in Sunnyvale, California. By working together and creatively using today's technologies, Dan and his team of efficiency experts are designing and building data centers that already meet or exceed EPA power efficiency recommendations for 2011.

Strategy for Data Center Power Efficiency

By Dan Hoffman

NetApp is completely committed to being an industry leader in data center energy conservation. We live this commitment by building highly efficient systems and consistently squeezing power inefficiencies out of our own data centers. We are fortunate to have a CEO who constantly challenges us to push the limits of what's possible.

This article summarizes our winning strategy for data center power efficiency. You can [read the full report](#) and take a [slideshow tour](#) of our Sunnyvale data center. You may also wish to read our report on [reducing data center power consumption](#).



The recent EPA report to Congress on data center power efficiency cites how we minimize power and cooling required to support data center operations. The EPA suggests that the Power Usage Effectiveness (PUE) rating be used as a measure of power usage throughout the data center. The current PUE benchmark is 2.0, which is equivalent to a power-to-cooling ratio of 1:1. At the time they go live in 2008, our new facilities in Sunnyvale, California, and Research Triangle Park (RTP) in Raleigh, North Carolina, will exceed the state-of-the-art scenario that the EPA recommends data centers achieve for power efficiency by 2011. The projected PUE for our new facilities is 1.2, which matches the EPA maximum achievable scenario for 2011.

We Collaborate and Make Power Efficiency a Priority Everywhere

The first step in getting serious about power efficiency required coordinating efforts across all areas of our company. We found that having a broader group at the table allows us to make better decisions on more effective use of our IT systems and facilities. When everyone understands the total cost of business systems and considers both power and cooling loads, we are able to drive greater efficiency while supporting aggressive business growth.

We regularly collaborate with our counterparts in the IT organization, NetApp workplace resources, and outside experts. In planning the recent retrofit of our Sunnyvale data center, we looked to an outside consulting company and to our local energy provider, PG&E, for help. The help offered by PG&E was invaluable. Not only were the professionals there able to offer recommendations, they also ended up financing important retrofit improvements through their generous rebate program. Rebates and power savings combined provided return on our investment in less than a year.

RELATED INFORMATION

- [Ten Techniques for Improving Data Center Power Efficiency \(PDF\)](#)
- [Data Center Power Efficiency from the Inside Out](#)
- [Reducing Data Center Power Consumption \(PDF\)](#)
- [EPA Data Center Report](#)
- [NetApp IT Storage Consolidation](#)
- [NetApp Helps Government Go Green](#)
- [eWeek eVideo "Scaling the Data Center"](#)

EPA Best Practices

In August 2007, the Environmental Protection Agency submitted a report to Congress regarding the need for higher power efficiency standards for data center systems. The agency noted that data centers consume roughly 1.5% of the total electricity consumed in the United States and projected that this consumption would double in the next five years.

In its report, the EPA constructed a set of scenarios that set objectives for improving the energy efficiency of data centers by 2011. The state-of-the-art scenario exploits the maximum efficiency potential for today's technology to reduce electricity use by up to 55% compared to current trends.

The best-practice scenario uses today's technologies to reduce electricity use by up to 45%. The improved operational management scenario uses only low-cost

We Attack the Problem from the Inside Out

A data center is a complex operation with lots of moving parts, any one of which can affect the power equation. A key aspect of our strategy to eliminate power inefficiency is to attack inefficiency on all fronts—starting from the IT systems and layout of the data center, moving to the supporting infrastructure that surrounds the data center, moving to the exterior walls of the building, and, finally, outside, where power generation happens.

Virtualize and consolidate IT systems. We started with storage consolidation—the foundation for reaping big benefits from server virtualization. By decommissioning 50 storage systems and consolidating onto 10 of the latest storage systems running Data ONTAP® 7G, we reduced our storage footprint from 25 to six racks. Equipment that drew 329kW was replaced with equipment that draws just 69kW. In the process, we eliminated 94 tons of air conditioning and reduced electricity costs to power those systems by \$60,000 per year. A major server virtualization project is currently under way to help provide additional savings in utilization, floor space, power, and cooling. Stay tuned for those results in future Tech OnTap articles.

Manage the data. By carefully analyzing business data, we discovered that 50% of the data we were storing could be eliminated.

Work with physics in data center layout. Using the laws of physics, rather than fighting them, is another way to save power. To begin, we alternate hot and cold aisles by placing equipment in racks front to front and back to back. Because cold air sinks, a blanket of cold air drops from above the front of the machines. This cold air is drawn into the machines and exits out the back as hot air. The hot air naturally rises to vents in the ceiling, where it exits the data center. We save a lot of energy by cooling only the systems (rather than the entire data center) and by isolating the hot air to keep it from being drawn back into the systems. Blanking panels in unused rack space, vinyl curtains, and physical barriers help further isolate the hot air.

Minimize electrical conversion losses. To minimize electrical conversion losses, we use a high-efficiency flywheel UPS. The flywheel UPS is more environmentally friendly, requires less maintenance, and is more efficient than a battery UPS. A battery UPS is 85% to 94% efficient, while the flywheel UPS is 97.7% efficient.

Exploit free cooling. "Economizer cooling" is an industry term for a cooling system that uses outside air for free cooling. Instead of recirculating air inside the data center, we bring in cool outside air and exhaust the hot air from the data center. This works only when outside air is below established temperature set points. In Sunnyvale, we use free or mixed cooling 68% of the time over the course of a year. In our new labs we will monitor humidity and temperature and calculate our use of outside air based on which process is more efficient.

Match fan speeds to actual cooling needs. Our variable frequency drives automatically reduce air handler fan speeds to match the airflow needs of a particular row as determined by rack sensors. The energy saved by reducing fan speed is exponential: a 50% reduction in fan speed yields an 87% reduction in power consumption.

Use heat from on-site electricity generation. To protect ourselves from spikes in energy costs as energy demand peaks, we depend on on-site equipment that uses natural gas to generate electricity. This combined heat and power system (CHP) enables us to produce chilled water without purchasing electricity. The heat generated by producing electricity is put to work as an energy source for our adsorption water chiller equipment. Additionally, our water-cooled electric chillers are more efficient than air-cooled chillers.

We Challenge Old Assumptions and Test New Hypotheses

Assumption 1: Facilities management and IT strategy are discrete functions.

At NetApp, we discovered years ago that teaming up results in more efficient systems. Having each group aware of the needs and challenges of others makes it possible for facilities to quickly respond to the changing growth demands from IT. At the same time, IT is aware of facilities challenges—floor space, power, and cooling—and can balance these concerns against the growth needs of the business.

measures to reduce electricity use by 20% relative to current trends.

[Read the report](#) to get the full details.

Power Usage Effectiveness (PUE)

The EPA uses the Power Usage Effectiveness (PUE) ratio as the metric for data center efficiency. $PUE = \text{total facility power} / \text{IT equipment power}$. The lower the PUE, the greater the efficiency. The current PUE benchmark is 2.0. The EPA state-of-the-art scenario for 2011 cites a PUE of 1.45 and a maximum achievable PUE of 1.2. In 2008, our new NetApp facilities will meet the 2011 predictions with an expected PUE of 1.2.

NetApp IT by the Numbers

Users: 8,500+ at 113 worldwide locations

Data centers: 6: Americas (3), Europe, India, Hong Kong

Servers: 1,700

Storage: 1.8+PB

Applications: 100+; six tier 1

2007 densities: 4kW–8kW per rack

2008 densities: 12kW per rack (new facilities)

Assumption 2: We've always done it this way.

Raised floors in data centers are artifacts of mainframe days. Our cold aisle/hot aisle layout made a raised floor unnecessary, which enabled us to put our power and cooling delivery overhead. And our newer data centers look more like a series of enclosed tunnels than one large room. This design provides the ultimate physical barrier between hot and cold air.

Assumption 3: High-tech problems require new technology to solve.

Power efficiency doesn't always require expensive high-tech solutions. Consider the vinyl curtain. You can hardly get more low-tech. By creating a physical barrier between hot and cold aisles, vinyl curtains save us one megawatt per year in Sunnyvale alone.

Assumption 4: Data centers have to be really cold.

Over the years we have gradually moved up the temperature in the data center from a starting point of 52°F. As we move the temperature up, we save in multiple ways: we need less cooling, we can get more hours of free cooling using outside air, and we can increase the water temperature in our chillers to maximize chiller efficiency.

We do realize that raising the data center temperature can reduce your margin of error in the event of a cooling system failure. The only way to protect against this is to continually monitor, measure, and tune your IT environment.

We Monitor, Measure, and Tune

Running at a higher temperature means a lower margin of error. In order to prevent disaster, we constantly monitor, measure, and tune a variety of environmental components in the data center, including temperature, cost of natural gas and electricity, and power loads. These may sound like minor things, but the accumulation of these little things makes a big impact. For example, we have sensors all over the data center to closely monitor temperatures and warn us of hot spots. We constantly monitor the cost of natural gas and the cost of electricity so we know when and how long to run our natural gas generators, and we constantly tune our cooling systems to cut down on the amount of energy wasted by overcooling our systems.

Taking Power Savings to the Limit

We are now taking what we've learned about maximizing power efficiency and putting it to work in our new data centers. In 2008, using technology available today, we will meet the EPA "what's possible" scenario for 2011. Twelve kilowatts will power the racks, and three kilowatts will cool them.

Here are a few of the forward-thinking techniques we're using:

- Pressurized cold rooms to eliminate the short cycling of air around the cabinet and reduce the energy required to power equipment fans
- Higher temperature set points to reduce overall cooling, provide more hours of free cooling using outside air, and maximize chiller efficiency
- Occupancy monitoring in hot aisles for automatic adjustment of lighting and cooling to conserve energy during the 90% of the time that the data center is unoccupied

Increasing power efficiency not only saves money but, according to the EPA, it can also help increase system reliability. For more information on how we are making huge strides in power efficiency using a combination of off-the-shelf products and lots of creative thinking, check out "[Ten Techniques for Improving Data Center Power Efficiency](#)."

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Efficiency Team Outside Sunnyvale Data Center

Left to Right.

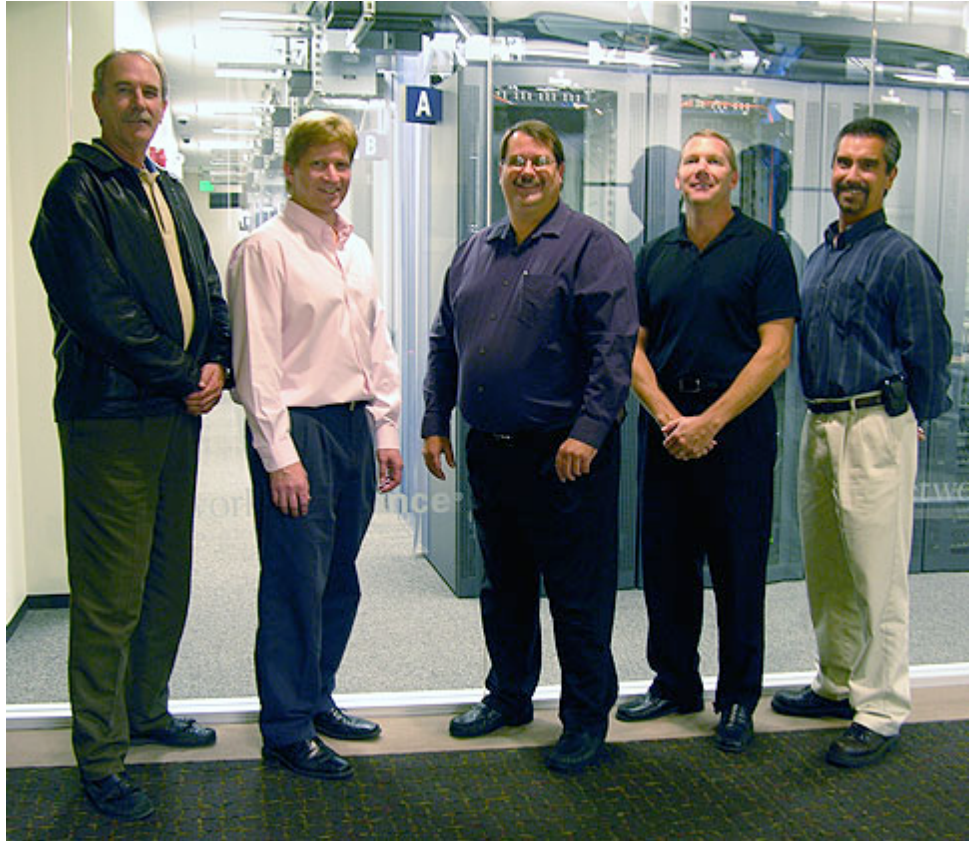
Rick Turner
Facilities MEP Manager

Dan Hoffman
Sr. Director of Facilities

Dave Shroyer
Facilities Controls Engineer

Randy Egger
Datacenter Capacity Manager

Ralph Renne
Site Operations Manager



< Previous | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Control Center

Our control center is staffed 24x7. Video cameras monitor all motion in the data center. On the center screen, you see Operations Manager, which we use to manage 80 storage systems and 1.8PB of data. On the right screen, an SMPT tool reports alerts and gives us an at-a-glance view of data center operations.



< Previous | 1 | **2** | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Hot Aisle/Cold Aisle

To keep one machine's exhaust from heating another, we place our racks front to front and back to back. This arrangement, called "hot aisle/cold aisle," has become a best practice for data center design because of its energy efficiency. Because the hot air is prevented from mixing with the cold air, less energy is needed to generate the cold air. For economizer operation, the cold air set point can be raised to allow more hours of "free cooling" from outside air.



< Previous | 1 | 2 | **3** | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Overhead Bus

Here you can see the overhead bus power delivery system. Each bus bar carries three phases of 220V power. With this system, our data center administrators can provision power in about five minutes—without risking a power disruption.



< Previous | 1 | 2 | 3 | **4** | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Strip Curtains

Vinyl curtains similar to the ones you see in meat lockers are used to contain the air in hot aisles. The material is specially designed to melt at lower temperatures so that in the event of a fire, the curtains drop to clear the way for the sprinkler system. The same vinyl is used to create a physical barrier around ducts and equipment.

These curtains alone save NetApp one million KWh of energy per year.



< Previous | 1 | 2 | 3 | 4 | **5** | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Blanking Panels

Low-tech blanking panels are used on the fronts of empty racks to block cold air from being pulled through empty racks.



< Previous | 1 | 2 | 3 | 4 | 5 | **6** | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

UPS System

Two 720kW rotary UPSs provide filtered, reliable backup power to critical equipment for up to 25 seconds or until diesel generators come online.

These UPS systems are 97.7% power efficient. As a side benefit, we have found the system is much more reliable than battery UPS systems, and there's no toxic battery waste.



< Previous | 1 | 2 | 3 | 4 | 5 | 6 | **7** | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Louvers on Outside Wall

When the temperature outside is lower than established set points, dampers behind the outside wall automatically open, allowing outside air to flow through the cooling system. When the temperature of the outside air is too hot to provide any free cooling, the dampers automatically close.



< Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | **8** | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Automatic Fan Speed Controls

Our variable-frequency drives automatically adjust air handler fan speeds to match the cooling needs of each aisle. Temperature sensors are in each aisle. The control system modulates the damper in each aisle to maintain the temperature set point.



< Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | **9** | 10 | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Economizer Air Handlers

Instead of recirculating air inside the data center, our economizer air handlers distribute cool outside air in the data center and exhaust the hot air through the building louvers. In Sunnyvale, we use outside air for free cooling on average about 68% of the year. We're in the process of gradually raising the temperature in our data center. By the time we finish, we expect to be able to use full or partial free cooling 97% of the time.



< Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | **10** | 11 | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Cogeneration Natural Gas Generators

Three 375kW gas-fired generators provide supplemental power during peak demand.

The heat produced by the generators is not wasted; it powers the adsorption water chiller equipment. By using the otherwise-wasted heat, our cogeneration is more efficient than that of most electrical power generation plants.



< Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | **11** | 12 | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Adsorption Chillers

The waste heat that is generated from producing energy with our natural gas cogenerators is put to work as an energy source for our adsorption water chiller equipment. The adsorption chillers use silica gel instead of refrigerant and are not ozone depleting. They are also very tolerant of the low-grade hot water coming from the generators. The adsorbers use the waste heat from the generators to provide "free cooling" with virtually no electricity used in the process.



< Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | **12** | 13 | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Mechanical Chillers

We use 250-ton rotary electric chillers that are water-cooled because they are more efficient than air-cooled chillers. The water-cooled chillers operate with our diesel generators in the event of a power outage. We also have an on-site well to provide unlimited make-up water to the cooling towers in the event of a main water line break due to an earthquake or other disaster.



< Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | **13** | 14 | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Main Electrical Board

The control board for the cogeneration provides parallel connection to the utility service for the data center.



< Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | **14** | 15 | Next >

© 2007 Network Appliance, Inc. All rights reserved.

Data Center Power Efficiency from the Inside Out

A guided tour of the NetApp Sunnyvale data center.

Diesel Generators

Two paralleled 2MW diesel generators can supply emergency power within 12 seconds of a utility outage.



< Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | **15** | Next >

© 2007 Network Appliance, Inc. All rights reserved. Specifications subject to change without notice. NetApp and the Network Appliance logo are registered trademarks and Network Appliance is a trademark of Network Appliance, Inc. in the U.S. and other countries. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.