

HIGHLIGHTS

- » Core NetApp DNA
- » A Thin Provisioning Case Study
- » Favorite Resources:
 - NetApp Visio Stencil Library
 - Tech Training: Intro to NetApp SANs
 - Report: Reducing Power Consumption
 - Webcast: Better Backup/Recovery

How NetApp IT Achieved 60% Utilization While Saving 41,184 kWh per Month

Stacey Rosenberry, Infrastructure Project Manager, NetApp

This detailed case study shows how an IT team cut its storage footprint from 25 to six racks and eliminated 94 tons of air conditioning.

Project Scope:

- 109 applications
- 343 servers
- 5,000 mounts

[More »](#)

BLOGGING WITH DAVE

Dave Hitz, NetApp Founder and EVP

"Values are a tool that employees can use to beat up management."

» [Dave's Blog](#)

DRILL DOWN

- » Reducing Power Consumption via Efficient Storage «early access»
An 8-point strategy to subtract machines and disks by using storage more efficiently.
- » A Better Way to Do Backup/Recovery
In this Webcast, hear how Agilent reduced backup windows by 98% (plus submit questions live!).
- » Results: Tech OnTap Annual Survey
Thanks for your feedback and ideas!

TIPS FROM THE TRENCHES

10 Steps to Secure NetApp FC SANs

Nick Triantos, Global SAN Systems Engineer, NetApp

FC SANs are vulnerable to WWNN spoofing, e-port replication, server viruses, and more. Are you 10 for 10 with these security precautions?

**10% off
NetApp
SAN
Training**

[More »](#)

[NetApp SAN Quick Reference Guide »](#)

[Tech Training: Intro to NetApp SANs »](#)

Technical Case Study:

Thin Provisioning for Disk-to-Disk Backup

Matthew Taylor, Professional Services Engineer, NetApp

See how a backup services provider boosted storage utilization over 70% by implementing thin provisioning on secondary storage. This detailed case study includes:

- Volume configuration
- Safety measures
- Monitoring practices

[More »](#)

ENGINEERING TALK

The Core NetApp DNA

Bruce Moxon, Sr. Director of Strategic Technology, NetApp

The WAFL file system, RAID-4, NVRAM, and a unique approach to snapshots represent the basic building blocks of NetApp technology. Learn how NetApp supports:

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

[More »](#)

ADMIN RESOURCES

» Guided Tour: NetApp Visio Library

Photorealistic accuracy, 3D shapes, dynamic "smart" shapes, PowerPoint icons, and more.



» Demo: SnapManager® for Oracle®

See how NetApp software integrates with Oracle ASM and enables the creation of entire database clones in minutes.

» Tech Training: Intro to NetApp SANs

See how NetApp storage relates to LUNs and more (plus save 10% on SAN training!).

If NetApp products were beers, what would you call them?

Aggregate Ale

Old Fiberslosh

Storage Security Stout

A-SIS Light

iSCSI Pale Ale

Colt 6070 Malt Liquor

[Submit your own!](#)



FEEDBACK



Stacey Rosenberry, Gary Garcia, Devinder Singh
NetApp IT Team

The storage consolidation project described in this article was a team effort. Stacey Rosenberry was the project manager, Gary Garcia was the project sponsor, and Devinder Singh served as IT architect. The authors have a combined 38 years of IT management experience. The project team is grateful to all those across the company who helped out. The initial team included John Lavrich, Victor Ifediora, and Sudeep Mullick, and the project could not have been accomplished without the support of the NetApp DBA and server teams and the cooperation of the application owners.

A Case Study: How NetApp IT Achieved 60% Utilization While Saving 40,000 kWh per Month

By Stacey Rosenberry, Gary Garcia, and Devinder Singh

In March, NetApp will publish a new white paper in the Vision series. As a Tech OnTap member, you're invited to enjoy early access to the report, [Reducing Data Center Power Consumption through Efficient Storage](#).

In 2006, NetApp IT undertook a project to increase utilization and upgrade hardware. This project required migrating from old, inefficiently used storage systems to new, more scalable systems. A significant benefit of the migration was the adoption of NetApp Data ONTAP® 7G and FlexVol™ technology.

This consolidation yielded significant results:

- Storage utilization increased from less than 40% to an average of 60%.
- Storage footprint reduced from 24.83 racks to 5.48.
- 50 storage systems replaced with 10.
- Direct power consumption decreased by 41,184 kWh per month.
- \$59,305 in annual electricity costs eliminated.
- Substantial capacity and performance gains.

This article explains the challenges that NetApp faced, the different phases of the consolidation, and key results. For details of exactly how we achieved (and calculated!) some of these results, see the sidebar.

The Challenge: Low Storage Utilization and Inefficient, Aging Hardware

Like many companies, Network Appliance has experienced rapid, sustained growth in recent years. With a 30% annual growth rate, simply adding more disks to our installed storage systems was not a viable long-term solution. The NetApp IT team was experiencing challenges in three key areas:

- **Low storage utilization.** Overall storage

Sneak Preview [Reducing Data Center Power Consumption through Efficient Storage](#)



NetApp is publishing a brand new vision white paper on reducing power consumption; available in March.

As a Tech OnTap member, you get a [sneak preview!](#)

RELATED INFORMATION

- [Sneak Peek: Reducing Data Center Power Consumption \(pdf\)](#)
- [Maximizing Storage Utilization \(pdf\)](#)
- [Technical Case Study: William Beaumont Hospitals \(pdf\)](#)

An Eight-Point Strategy for Reducing Storage Power Consumption

Powering the data center has quickly become one of the top issues that enterprises face today. Customers are asking how NetApp can help them reduce power consumption.

The NetApp approach to fighting rapidly growing power consumption is simple: subtract machines and disks from the power equation by using storage more efficiently.

This white paper outlines the NetApp eight-point strategy for reducing storage power consumption.

Read the [white paper \(pdf\)](#).

Calculating Power Consumption Savings

In total, the storage equipment that we decommissioned drew a maximum of 1631 amps, or 329kW, and was replaced with equipment drawing a maximum of 331 amps, or 69kW.

Based purely on faceplate value, therefore, this project eliminated 260 kWh. If this data center were located in a co-location facility, power

Aging hardware. This project focused on a variety of older hardware, including 34 F760s, 12 F820s and F840s, and 4 F880s. These systems were running older versions of the Data ONTAP operating system, which did not allow the team to take advantage of advanced features such as [FlexVol](#) technology. These older systems also use lower capacity drives with lower overall storage density, resulting in a storage environment with a large number of management complexity.

- When we started the upgrade project, we realized that this was not just an infrastructure process; bringing our business applications up to modern best practices also required that we rationalize the network topology, the data storage layouts and application code. Our project methodology was adapted to integrate with each application team, using planned software release windows opportunistically. Although we primarily set out to tackle our storage issues, it was impossible to ignore the rest of the environment.

- This storage environment supports a wide variety of critical business applications used by more than 20 business groups. One thing that worked to our advantage was that rather than distributing business-critical applications across multiple worldwide data centers, in most cases NetApp was already using a single global instance of applications, reducing complexity versus enterprises that have widely distributed applications.

- Naturally, the applications were spread across an even larger number of servers. The impact of the storage migration to each server had to be assessed, and each server had to be migrated to the new storage environment. The difficult part was not the server to storage relationship, but rather the relationships from a shared storage infrastructure to the application set. In effect, the migration was many application migrations, the servers to storage relationships were simply the context.

- Network Appliance had adopted a segmented network strategy, but legacy systems still depended primarily on one monolithic flat network that mixed development and production and exposed applications unnecessarily to network "weather". This project gave us an excellent opportunity to bring legacy systems into best practices.

- The upfront coordination and support of NetApp application developers and storage administrators was a must. We also had to coordinate the efforts of the DBA, UNIX® server, and Windows® server teams. Without buy-in from management to have these resources available, this project wouldn't have been completed.

[illegible]

Tech OnTap February 2007 | Page 3

- Our initial discovery indicated that we needed to consider 109 different applications. Each application had at least two environments (development and production), while some tier 1 application had as many as eight discrete environments.
- These applications were utilizing 343 servers. By talking to application owners, we found that 148 of these servers would not require migration and that 18 could be decommissioned. This left 177 servers whose data would need to be migrated to consolidated storage.
- Application storage was being provided by 50 separate storage systems with 53.6TB of stored data on 331 volumes. We discovered just under 5,161 mounts to these servers. In many cases, information was hard-coded and would need to be changed by each application team before we could proceed.

Best Practice for Coding NFS Mount Point References

- Bad: coded mount points (i.e. filename:/vol/app/qtree/directory)
- Good: coded paths (i.e. /netapp/oracle, or /netapp/gnu)

Phase II: Analysis

Based on the audit, NetApp IT decided to implement the following changes:

1. **Decommission 50 storage systems** and replace them with 10 of the latest model storage systems (at that time, the FAS980c) running Data ONTAP 7G.
2. Host the new storage systems in **segmented networks** so that performance could be better managed between applications.
3. **Migrate existing servers** to the new network infrastructure.
4. **Migrate 46 applications.** We decided 44 applications were already compliant with storage standards and learned that 19 could be decommissioned.
5. **Convert all mounts to standardized references;** eliminate all references to specific storage systems.

This sounds simple enough, but it represents a significant amount of change with a lot of dependencies. As Dave Robbins, senior director of NetApp Global Infrastructure, pointed out, "NetApp IT may own the plumbing, but the application folks own the furniture and ultimately we can't screw up the house during the remodel."

Phase III: Implementation

The project began with an intensive manual process of cleaning up the data. Every data set had to be reviewed. We had developed scripts that allowed us to do an inventory of mount points—where they are connected, etc.—but ultimately each mount had to be scrutinized by someone from the responsible application team, and each team had to decide what to keep, what to archive, and what to delete. Programmers also had to go back and fix any hard-coded mounts and other dependencies that would break during the migration.

Next, we installed the new storage systems and configured new networks utilizing segmented VLANs to isolate application traffic. With those tasks complete, data migration could begin. We worked through the applications one at a time. For each application we established a migration team and developed a move plan. Two to four application projects were run concurrently. Actual data movement was carried out using either NDMPcopy or NetApp SnapMirror® replication software. Once an application was migrated, we made old volumes obsolete and decommissioned old storage systems.

Project Results

The storage consolidation phase of this technology refresh has provided a broad range of benefits addressing the storage challenges previously described above.

Challenge #1: Low storage utilization

Result: An average of 60% storage utilization

- **Disk utilization increased from about 40% to more than 60%.** This was a direct result of the move to Data ONTAP 7G and FlexVol. Using flexible

In May, NetApp user Ben Rockwood provided his own overview of how the NetApp Data ONTAP operating system manages data on disk and how that data is accessed from client systems.

Highlighted technologies include:

- RAID, RAID-DP, and traditional volumes
- Aggregates and FlexVol volumes
- Snapshot functionality and FlexClone technology
- LUN creation and masking

Read More:

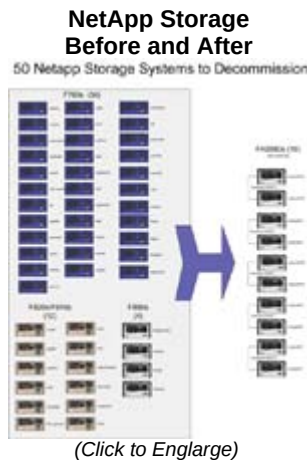
- [A User Perspective on 7G](#)
- [NetApp Technical Report: Introduction to Data ONTAP 7G \(pdf\)](#)

volumes, we have been able to spread application volumes across a large number of spindles for performance without sacrificing disk space. Increased utilization means that we need less total disk capacity, decreasing power consumption and cooling requirements and simplifying management. For the Cognos application highlighted previously, for example, utilization jumped from an average of 28% across 8 storage systems (a high of 80% and low of 4%) to an average of 85%.

Challenge #2: Aging, inefficient hardware

Result: Significant gains in capacity, performance, flexibility, reliability, and ease of management

- **Increased capacity and performance.** Although in the short term we reduced our storage requirement by improving utilization, this upgrade also positions NetApp to quickly expand storage capacity in the data center as necessary. Replacement of older disks with 144GB disks substantially increases the capacity of each disk shelf. Each of the new systems has a maximum capacity of 64TB, meaning that the 10 storage systems deployed can support up to 640TB. These 10 storage systems also offer significantly more performance and capability than the 50 systems they replaced.
- **Increased operation flexibility.** The move to consolidated storage on Data ONTAP 7G makes it much easier to add capacity (and less expensive as a result of better utilization). With FlexVol volumes, we can easily add new volumes or grow or shrink existing volumes to meet changing demands.
- **Increased stability and reliability.** All new storage systems are clustered for improved data availability, and all RAID groups utilize RAID-DP™ for greater protection against disk failure. Using diagonal parity, RAID-DP can recover from two disk failures in the same RAID group, yet offers the same performance as NetApp RAID 4.
- **Simplified management by replacing 50 storage systems with 10.** Now we have only 10 storage systems to manage, and we took care to rationalize volume names, mounts, and exports while eliminating hard-coded dependencies to ensure smoother operations going forward.



Challenge #3: Space, cooling, and power constraints.

Result: Reduced storage footprint to under 6 racks and cut annual power costs by \$60,000.

- **Substantially reduced data center footprint.** As shown in the following table, through this consolidation we've been able to reduce our storage footprint from 24.83 standard 47U foot racks to 5.48 racks.
- **Reduced power consumption and electricity costs.** In total, the storage equipment that we decommissioned drew a maximum of 1631 amps, or 329kW, and was replaced with equipment drawing a maximum of 331 amps, or 69kW. This resulted in an electricity savings estimated at \$59,305 annually (see the sidebar for details). Additionally, the resulting decrease in heat load works out to 93.549 tons of air conditioning.

	Original	After Consolidation
Rack Space	24.83	5.48
Disk Utilization	<40%	>60%
Direct Power Usage	329kW (Max)	69kW (Max)
Estimated Annual Power Savings	-	\$60,000

Finally, as part of this project, the team reorganized the NetApp network infrastructure. The segmented network architecture allows us to isolate application traffic using VLANs for better, more predictable performance and improved security.

[Comment on this article](#)



NICK TRIANTOS

Global SAN systems engineer, NetApp

A member of the elite Global Systems Engineering group, Nick helps top enterprise companies solve their toughest technical challenges. Nick has been in systems or support engineering roles, including positions at HP as an account support engineer (Server group) and presales technical consultant (Storage group) for nearly 16 years. Nick maintains a blog and has authored multiple Tech OnTap articles.

Ten Steps to Secure Fibre Channel SANs

By Nick Triantos

For a long time, Fibre Channel SANs have been considered secure primarily because they are deployed as closed, isolated networks within the data center. While physical network isolation offers a level of security, a breach on any host connecting to the fabric could allow unauthorized access to the SAN.

Several well-known techniques, such as WWNN spoofing and E-Port replication, exist for gaining unauthorized access to storage. Also, servers infected by viruses can be exploited and serve as gateways to the fabric.

SAN security always begins at the OS level and progresses to the other elements, such as applications, switches, disk arrays, and management stations. Here are 10 tips to help enhance your FC SAN security:

1. Hard Zoning

Zoning has always provided good security against threats by logically isolating devices in a fabric. Zones are enforced either in hardware or software, depending on the chosen zoning scheme.

With some switch manufacturers, hardware zoning is enforced in the Nameserver *and* at the ASIC level. When *either* the WWN or the port number is used in the zone, that zone is *hardware* enforced. If the WWN is mixed with some elements of the port number in the zone, that zone is *software* enforced. To prevent zone-hopping attacks, ensure that zones are using either the switch port number or a device's WWPN instead of the WWNN. Given that the WWNN of a device can easily be changed, it is too simple to spoof it and become a member of a zone.

2. LUN Masking

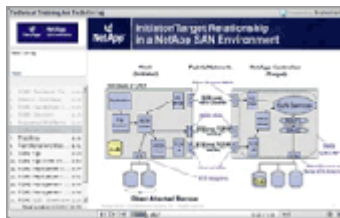
LUN masking can be performed at various points in the SAN (array, switch, host). The safest and most secure point to implement LUN masking is at the closest point to the source device—the disk array. Furthermore, it should be implemented using WWPNS instead of WWNNs, again because the latter

NetApp University Technical Training Introduction to NetApp SANs

This 45-minute technical training explains how NetApp storage systems work in a SAN environment.

Includes a discussion on the relationship between storage and LUNs, space management, and tools that facilitate host integration.

Watch now!



Sign up for **Data ONTAP® SAN Admin Basics, 7.2** before April 15 and save 10%
(NOW login access required; discount applied automatically during registration)

RELATED INFORMATION

- [FREE: 45-Minute Technical Training: Introduction to NetApp SANs](#)
- [NetApp SAN Reference Guide \(pdf\)](#)
- Previous articles by Nick:
 - [iSCSI vs. FC SAN](#)
 - [SAN Implementation Tips](#)
 - [Dynamic Queue Mgmt.](#)
 - [NetApp Technical Reports on SAN](#)

NetApp SAN Environments - Quick Reference Guide -

Originally developed for NetApp employees and channel partners, this two-page document summarizes the components of a NetApp SAN environment.

Includes specs and feature summaries for:

- NetApp SAN hardware platforms
- NetApp SAN software
- NetApp SAN service offerings
- SAN hosts supported by NetApp
- SAN switches supported by NetApp

	FC	FC	FC	FC	FC	FC	FC
Max. Host Capacity (TB)	514	420	100	250	186	84	15
Max. # Hosts	1,000	500	672	500	320	100	50
Host Type	FC	FC	FC	FC	FC	FC	FC
Max. FC Ports	40	40	30	20	20	20	4
Max. # LUNs	2048	2048	2048	2048	1024	1024	1024

[Click to Enlarge](#)

[Read the Quick Reference Guide.](#)

can be easily spoofed.

3. Port Binding

Port binding allows access to the fabric via a specific switch port based on the WWNN or WWPN of the connecting device. No other device can connect to that specific port. The switch itself maintains a database listing all devices that are bound to specific switch ports. Again, the use of WWPN is recommended.

4. Port Type Locking

Switch ports by default are of type Generic (G-Port). That means that depending on what's connected to the port, the port itself may assume several possible modes of operation (E-Port, F-Port, FL-Port, etc.). Port type locking allows you to restrict a switch port to a specific mode of operation, thereby limiting it to a certain task. This prevents the possibility of an unauthorized switch joining the fabric.

5. Logical Partitions

Logical or virtual switch partitions provide further isolation above zoning at both the protocol processing layer and the management layer. By default, routing of traffic between devices in different virtual switch partitions is prohibited. Additionally, logical or virtual partitions provide multiple instances of fabric services on a per virtual partition basis, increasing scalability and availability.

6. Unused Switch Ports

Disable unused switch ports. It is not uncommon to find all switch ports enabled within a fabric, even though only a fraction of them are in use. With these ports enabled, an unauthorized device can enter the fabric and potentially disrupt it or access unauthorized data.

7. SPAN Ports

A Switched Port Analyzer (SPAN) allows copies of frames destined for a device through an FC port to be forwarded to a SPAN destination port. This is used mainly for low-level troubleshooting and allows for the capture and analysis of each frame. However, this is a security risk since traffic can be captured on any port. Use role-based access on the switch to restrict unauthorized users from enabling SPAN.

8. Secure Shell (SSH) and HTTPS (SSL)

The problem with using telnet as a way to access and manage an FC switch or a disk array is that it sends the login name and password in clear text. Using tools such as Ethereal, one can easily obtain this information and gain access to the switch. Secure Shell supports various encryption algorithms (3DES, Blowfish, AES, and Arcfour) and provides strong authentication, thus securing access to storage resources.

HTTP also transmits authentication information in an HTML format subject to easy capture by sniffing the LAN connection. Both the login and password are Base-64 encoded as they traverse the network. A Base-64 decoder (which can easily be found on the Internet) can be used to decode the authorization string and obtain both the login and password. HTTPS encrypts all traffic to and from the target device (e.g., FC switch, array, or management station), avoiding this security loophole.

9. Fibre Channel Security (DH-CHAP) and SNMPv3

DH-CHAP is a secure authentication protocol that supports both the MD-5 and SHA-1 algorithms, providing switch-to-switch and host-to-switch authentication. Today, pretty much every switch and HBA vendor supports DH-CHAP. Switch-to-switch authentication is important for switches and switch links that connect fabrics across distributed data centers. It's also important to establish switch-to-host (HBA) authentication given that unauthorized access to data typically occurs at the host. SNMPv3 (Simple Network Management Protocol version 3) is an application layer protocol used by network management systems to monitor or manage devices in the network. Previous SNMP versions lacked authentication capabilities and potentially result in a variety of security threats. SNMPv3 provides authentication and integrity, as well as encryption. SNMPv3 traffic is encrypted with DES and carries an MD5 HMAC or an SHA HMAC algorithm for authentication and integrity purposes. If SNMP access is not needed, then it should be disabled.

10. Passwords and Event Logging

While it seems intuitive to change the default device password as provided by the manufacturer, it's amazing how many devices—FC switches in particular—are put in production using default passwords. It's very important to implement strong password policies for all users and manage access permissions by role rather than by user.

SearchStorage All-in-One Research Guide: SAN

If you're past storage basics, this guide will help you through it all.

- [SAN Management](#)—Detailed information on RAID configuration, provisioning, performance capacity management, and troubleshooting.
- [SAN Connectivity](#)—Find the latest information on Ethernet, IP, Fibre Channel, and iSCSI.
- [SAN Switches](#)—Learn about the characteristics of some of the most popular switch classes: blade, director, and intelligent.

Consolidating Your FC SAN

In this TechTalk Webcast, Eric Tomasi, VP of Infrastructure for Folksamerica Reinsurance Company, explains how his team leveraged a NetApp FC SAN solution to achieved 80% storage growth without increasing IT headcount.

Today the company can complete a full remote recovery in less than 4 hours.

[Watch the Webcast](#) to learn more.



Additionally, implement single-user sign-on, if possible, using RADIUS, since it supports dynamic passwords and challenge/response passwords that make it very difficult for password-generation algorithms or phantom hosts that try to spoof users into giving up their passwords. Also, ensure that all successful and unsuccessful events are logged on a centralized server and analyzed. Finally, ensure that all devices are time synchronized so that events across devices are easily correlated.

Be Proactive

The rapid adoption of the Internet has eroded the belief that FC SANs are safe from attacks, making storage security an important consideration for every enterprise. Proactively implementing strong security mechanisms is the only way to guard against attacks that could compromise sensitive data or disrupt operations.

Learn More!

Free 45-minute training course: Intro to NetApp SAN Environments

Includes a discussion on the relationship between storage and LUNs, space management, and tools that facilitate host integration.

NetApp SAN Quick Reference Guide

Originally developed for NetApp employees and channel partners, this two-page document summarizes the components of a NetApp SAN environment and provides detailed specifications and feature summaries.

Save 10% on Data ONTAP SAN Administration, Release 7.2

This instructor-led course is designed for those who provide support and administration for FC and IP SAN environments running the Data ONTAP® operating system. (NOW login access required; discount applied automatically during registration)

Comment on this Article



Matthew Taylor

Professional Services Engineer, NetApp Global Services

Matthew Taylor has been working with NetApp storage for more than seven years. Prior to joining NetApp, he worked as a Windows® and storage administrator for a large manufacturing company. Matt joined NetApp in July 2005 and since that time has worked on-site supporting the top enterprise account described in this article. During that time, he has helped the customer's multiple business units grow their NetApp storage environment from 30 systems to over 90.

Technical Case Study: Thin Provisioning for Disk-to-Disk Backup

By Matthew Taylor

With traditional storage provisioning, you rely on storage end users to identify their requirements and then allocate all the disk space they think they will need for each application upfront. Unfortunately, end users are notoriously unreliable at estimating requirements. If you allocate a 1TB volume for the new "killer app" that's guaranteed to be a success, nine times out of 10 when you go back and look a year later you'll find that it's only using half the allocated space (or less).

When it comes to disk-to-disk backup, provisioning can be even more complicated. You not only need an estimate of the growth in primary storage usage, you also need to know the rate of change in each volume. This case study looks at an innovative customer application of thin provisioning on secondary storage for disk-to-disk backup to overcome these uncertainties. Over the course of a year the customer increased primary storage capacity from 500TB to 900TB without needing any additional secondary storage. This was a direct result of tremendous increases in disk utilization through the use of thin provisioning.

This article provides an introduction to thin provisioning in a NetApp environment and documents a real-world implementation, including:

- Change rates
- Safety measures
- Volume configuration
- Volume settings
- Monitoring practices

Background: How NetApp Approaches Thin Provisioning

Thin provisioning addresses the limitations of the traditional approach to storage provisioning. Conceptually, it works the same way as insurance. A typical insurance company holds policies far in excess of what it can pay out at one time, but the number of actual claims in a given period never exceeds the company's working capital, and it stays in the black. Having a large enough and diverse enough pool of customers helps ensure that an insurance company creates a risk-sharing versus risk-taking environment.

Similarly, with thin provisioning a storage system presents more storage space to the servers connecting to it than it actually has available. Consider a storage system with 15TB of usable storage capacity. With thin provisioning, a storage administrator may map volumes of 0.5TB to each of 45 servers, making 22.5TB of storage visible to hosts.

Free space on the storage system serves as a buffer pool for all volumes. Physical storage space is allocated to each volume on demand as data is written, so if all 45 hosts use the space provisioned to them there would obviously be a problem. You have to monitor the storage

This level of
simplicity in
configuring thin

Tech OnTap February 2007 | Page 9

RELATED INFORMATION

- [Podcast: Smart Storage Allocation through Thin Provisioning](#)
- [Technical Report: Thin Provisioning for NetApp SANs \(pdf\)](#)
- [Introduction to Data ONTAP 7G \(pdf\)](#)
- [Dave's Blog: How Thin Provisioning Helps Admin Write Bad Checks](#)

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 it. 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.](#)

The Versatile Storage Platform

To truly appreciate the versatility of the NetApp architecture, it is important to view how the storage is managed and how the storage is

system and add capacity when needed, but instead of making capacity planning decisions and provisioning to meet the needs of each individual volume, you plan and provision for the needs of the entire storage system. This is easier, less prone to mistakes, and results in much more efficient storage utilization, so less storage is needed.

provisioning is unique to NetApp.

NetApp Data ONTAP® 7G with FlexVol® technology provides a built-in mechanism for enabling thin provisioning. By simply setting the guarantee parameter on each volume to an appropriate value, thin provisioning can be enabled without host or application customization. This level of simplicity in configuring thin provisioning is unique to NetApp.

When you create a volume on a NetApp system, you don't have to dedicate specific disk blocks to the volume. Instead, blocks are allocated on demand as data is written. In this way, multiple volumes share the same pool of free storage, and you don't have to guess up front which volumes will grow and by how much. You simply add more capacity when free storage gets low and grow a volume if more space is required. With FlexVol you don't pay a performance penalty for this approach. Even the smallest volumes utilize a large number of disks for optimal performance.

Case Study: Thin Provisioning for Disk-to-Disk Backup

The customer described in this case study is a large company that sells backup services to its internal customers with a fixed retention guarantee (normally 45 days). The customer utilizes NetApp storage systems for both primary and secondary storage. All systems are running Data ONTAP 7G.

Primary Storage and Application Environment

Characteristics of the primary storage requiring backup include:

- 72 NetApp storage systems with about 900TB of capacity.
- Retain seven days of Snapshot™ copies on local storage for quick restores.
- Databases range from 100GB to 6TB in size.
- Approximately 150 groups and services are served by this storage.

Oracle® Databases are the most critical and most volatile of the applications supported by this storage. Each database is considered independent; this means that it must be possible to back up and (more important) restore each one individually. Database turnover is often very low, but at times may reach a 100% rate of change because of people loading new information. The storage team has no control or visibility into what might occur on particular primary storage volumes, so the backup system has to adapt readily.

Secondary Storage and Disk-to-Disk Backup Environment

The secondary storage and backup environment consists of:

- Six NetApp NearStore® R200 storage systems using 320GB SATA disks with approximately 430TB of total raw capacity
- NetApp SnapVault® software:
- SnapVault starts with a baseline copy on secondary storage that mirrors the source volume or qtree. (A qtree is a subvolume that has its own quotas and permissions.)
- When a nightly backup is scheduled, SnapVault is used to create a Snapshot copy of the primary volume and transfers only the blocks that have changed to secondary storage. (For databases, in house scripts put the database in hot backup mode before creating a Snapshot copy.)
- Snapshot copies are maintained on secondary storage for a prescribed time so that data can be restored from any point in time.
- Approximately 800 qtrees are in SnapVault relationships.
- From 14 to 45 days' worth of SnapVault backups are retained for each qtree.

Why Thin Provisioning?

After about a year running this configuration, it became clear to the customer that differences in the change rate of different data sets were resulting in significant underutilization of the R200 systems. Utilization was only at 40%, and yet the IT team was always concerned about secondary storage space since it was almost fully allocated. Manually managing 800 separate qtrees was impractical and painful.

The IT team was initially considering the concept of thin provisioning for another storage project. When NetApp demonstrated thin provisioning to the company's

accessed as related to functionality. Too often storage vendors separate these two concepts, which creates overly specialized storage systems that eventually become isolated islands.

In May, NetApp user Ben Rockwood provided his own overview of how the NetApp Data ONTAP operating system manages data on disk and how that data is accessed from client systems.

Highlighted technologies include:

- RAID, RAID-DP™, and traditional volumes
- Aggregates and FlexVol volumes
- Snapshot functionality and FlexClone™ technology
- LUN creation and masking

Read more:

- [A User Perspective on Data ONTAP 7G](#)
- [Technical Report: Introduction to Data ONTAP 7G \(pdf\)](#)

A Quick Primer on NetApp Data Protection Software

NetApp customers have two potential alternatives for data protection: SnapMirror® or SnapVault software.

SnapMirror is replication software intended for disaster recovery solutions. The mirror is an exact replica of data on the primary storage that can be mounted read/write to recover from failure. If a backup is deleted on the source, it will go away on the mirror at the next replication.

SnapVault, in contrast, is intended for disk-to-disk backup. It retains all backup copies as they appeared at the time they were created on primary storage for a user-specified period of time. Secondary storage used by SnapVault cannot be mounted read/write. Backups must be recovered from secondary storage to the original or an alternative primary storage system in order to restart.

At a more technical level, SnapVault takes a point-in-time image based on qtrees, while SnapMirror copies an entire image at the level of a LUN inside a volume.

Get the details. Read the reports:

- [SnapMirror Best Practices Guide \(pdf\)](#)
- [Enabling Rapid Recovery with SnapVault \(pdf\)](#)
- [Data Protection for NetApp Storage Systems \(pdf\)](#)

storage administrators, however, the team recognized an opportunity to leverage this approach to solve its backup challenges. The team found thin provisioning more appealing for its backup environment because performance wasn't as big a concern—secondary storage was only occasionally accessed for restores—and it was possible to make changes to the backup environment as necessary (move qtrees to new aggregates and so on) without impacting production applications.

Converting to Thin Provisioning

Implementing thin provisioning was easy. The IT team simply made two adjustments to the volumes that housed the secondary qtrees for each SnapVault relationship:

- Changed volume guarantee setting to “none.”
- Sized each volume to match the size of the aggregate containing the volume.

With these changes, any volume can potentially grow to the full size of its aggregate, but no volume is guaranteed space. All volumes are free to grow as long as free space exists.

As a safety measure, the company created one “fully guaranteed” volume in each aggregate containing 20% of the total space. In normal operation this volume is not used but serves only as an emergency or backstop. If an aggregate were to fill unexpectedly, a storage administrator could release this space so operations can continue while rebalancing the distribution of qtrees between different aggregates.

The actual conversion process took some time because of the 800 qtrees in SnapVault relationships. To convert, they had to do a linear progression volume by volume and qtree by qtree. The company also used this as an opportunity to remap its qtree-to-volume relationships, which increased the total time for the conversion.

Changes to Monitoring Practices

A set of best practices was established that called for “administrative closing” of aggregates to new SnapVault secondary qtrees after the aggregate became 60% full and for outmigration of qtrees to other aggregates to begin at 85% full.

This actually reduced the number of qtree migrations between aggregates on secondary storage versus the previous traditional provisioning environment. The overly large space demands of the old model made free space a problem and required more frequent moves. The alerting and monitoring done by NetApp Operations Manager (formerly known as DataFabric[r] Manager, or DFM) was customized to account for the oversubscription of the aggregates and the need to provide a more appropriate “aggregate full” threshold. The company also changed from a policy of monitoring free space on volumes to monitoring free space on aggregates.

Result: 70% Utilization, No New Secondary Storage Despite 80% Increase in Primary Storage Capacity

This thin provisioning methodology has been in place for a year with no outages, and no aggregates have been filled. Before the migration started, the company was concerned with free space almost every day, but as the migration went forward, it continuously got back free space from formerly underutilized volumes. This free space made it possible to add new customers and services into the backup system without purchasing additional storage. Over the course of the last year, primary storage capacity has grown from 500TB to 900TB without requiring any additional secondary storage capacity. Before the switch to thin provisioning, the company had been considering adding an additional R200.

This particular data center was continuously pinched for floor space, power, and cooling, so this savings represents a significant benefit beyond the savings in capital outlays. The company has now been able to delay the purchase of any new secondary storage for a year as a result of thin provisioning and the increased efficiency it provides. Storage utilization went from less than 40% (due to mostly underutilized volumes) to closer to 70%.

Customer Recommendations

This customer doesn't hesitate to recommend the use of thin provisioning in a disk-to-disk backup environment. The company also uses thin provisioning for home directories on the production side of the house. According to company practice, each of 4,500 users has up to 1GB of network file storage as a home directory, which would require 4.5TB of total storage. Using thin provisioning, the company meets this requirement with only 600GB of actual disk storage.

Despite these successes, the customer is quick to point out that thin provisioning may not be appropriate for all uses. For OLTP applications, for instance, it is much harder to move data around without impacting the application should storage become critical, so important database applications are probably not a good choice for thin provisioning or should not be aggressively thin provisioned.



Bruce Moxon

Senior Director of Strategic Technology and Grid Guru, NetApp

Bruce Moxon works with enterprise customers deploying grid computing solutions. He brings more than 20 years of experience in scale-out computing architectures for both scientific and commercial applications and writes, speaks, and teaches extensively on the continuing evolution of grid computing. Bruce has architected and developed solutions for a number of high-throughput computing environments, including Perlegen Sciences' SNP discovery system, Bank of America's CRM and analytics systems, and NASA's Earth Observing System.

The Core NetApp DNA

This article draws significantly on core concepts described in TR-3002, [File System Design for an NFS File Server Appliance](#).

DNA. It's the building block of life—the macromolecule that encodes genes and governs the production of proteins, from which all cellular metabolism derives. It is truly the “molecular blueprint” that determines the properties of all organisms.

I often talk about the Network Appliance™ storage system architecture as having its own “DNA”—a core blueprint from which numerous key features derive and which continues to spawn new evolutionary “variations” that allow the architecture to adapt to environmental changes.

In NetApp storage architectures, this blueprint is based on the WAFL® file system, RAID 4, NetApp use of NVRAM, and a unique approach to Snapshot™ copies. These are the core building blocks that continue to define NetApp storage systems. And they continue to support the evolution of features that carry forth the core DNA—whether that be data protection and retention features (SnapMirror® and SnapVault®), compliance features (LockVault™), or efficient means of replicating working data sets for dev/test/QA environments (FlexClone™).

The Genes: WAFL, RAID 4, NVRAM, and Snapshot

At the core of the NetApp genetic blueprint are four key, interrelated technologies: WAFL, RAID 4, NVRAM, and Snapshot.

WAFL is the Write Anywhere File Layout, an approach to writing data to disk locations that minimizes the historic RAID write penalty. By keeping file system metadata (inodes, block maps, and inode maps) in files, WAFL is able to write file system metadata blocks anywhere on the disk. This approach in turn allows multiple writes to be “gathered” and scheduled to the same RAID stripe—eliminating the traditional read-modify-write penalty prevalent in parity-based RAID schemes.

In the case of WAFL, this stripe-at-a-time write approach makes RAID 4 a viable (and even preferred) parity scheme. At the time of its design, the common wisdom was that RAID 4 (which uses a dedicated parity drive) presented a bottleneck for write operations because writes that would otherwise be spread across the data drives would all have to update the single parity drive in the RAID group. WAFL and full-stripe writes, however, eliminate the potential bottleneck and, in fact, provide a highly optimized write path.

This stripe-at-a-time approach to writes also required that the system provide a means of reliably buffering write requests before they are written (en masse) to disk. Nonvolatile RAM allows the system to reliably log writes and quickly acknowledge those writes back to clients.

The final core contribution to the NetApp DNA is the implementation of Snapshot technology, which provides an efficient, point-in-time, consistent view of the file system. Figure 1a presents a simplified view of the WAFL file system (leaving out internal inode and indirect block structures). Figure 1b shows how WAFL creates a new Snapshot copy by simply duplicating the root inode. Both the original root inode and the

RELATED INFORMATION

- Previous articles by Bruce:
 - [Fueling the Enterprise Grid](#)
 - [Improving Database Agility](#)
- [TR-3002: File System Design \(pdf\)](#)
- [TR 3001: A Storage Networking Appliance \(pdf\)](#)
- [FlexVol® and FlexClone Demo](#)

A Storage Networking Appliance

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:

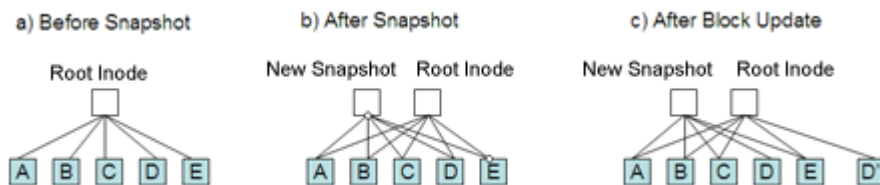
- The WAFL file system
- Snapshot copies
- Consistency points and NVRAM
- FlexVol and FlexClone technology
- RAID and RAID-DP™

If you only read one paper about NetApp technology, read [A Storage Networking Appliance \(pdf\)](#)

Architecting Storage for Resiliency

RAID-DP significantly increases data protection, with zero to minimal impact on capacity utilization and

Snapshot copy then point to the same blocks on disk (same “view” of the file system). Figure 1c shows what happens when one of the baseline file system blocks (block D) is modified by a user process. Only the new data (single write) need be written to disk. This write and any required modifications to intermediate nodes (inode blocks, indirect block maps) are logged into NVRAM, where they can be gathered and coalesced to optimize updates of those intermediate nodes.



The underlying layout, coupled with the episodic, multistripe write approach, ensures that NetApp Snapshot technology is extremely space- and resource-efficient. Effectively, only changed blocks (changes to the baseline file system) are written to disk. The result is that many Snapshot copies can be maintained as efficiently as one; more efficient Snapshot copies allow organizations to create Snapshot copies more frequently, which ensures faster and more up-to-date file or file system recovery.

In addition to providing read-only, point-in-time “versions” of the user file system, Snapshot copies are also used to create periodic “consistency points” within the file system that minimize recovery time in the event of power loss or system failure. These consistency points are taken every few seconds and together with NVRAM-journalled writes ensure rapid recovery of a consistent file system without the need for extensive consistency checks.

Evolution

These are the core building blocks of NetApp storage systems. Over the years, these core technology “genes” have **recombined** in a number of ways to deliver more and more capable storage systems—the analog of genetic evolution.

Block Storage Protocols

Initial NetApp storage systems were NFS appliances. Over the years, the same core architecture has been extended to support multiple protocols—CIFS initially, and then block-based protocols (Fibre Channel and iSCSI). Block protocols expose LUNs, which are special WAFL containers (“files”) that exhibit block device characteristics. They inherit the rich lineage of WAFL—including space- and resource-efficient Snapshot copies and clones.

Cluster Failover

The core NVRAM-based write journaling mechanism has been extended in conjunction with controller pairing to provide HA failover capabilities. In these clustered configurations, two controllers are “cross-connected” to each others’ disks, and NVRAM writes are mirrored over an InfiniBand® cable to the partner controller’s NVRAM, ensuring redundant journaling in case of controller failure.

SnapVault, SnapMirror

Snapshot copies are “on-box” point-in-time versions of a file system or LUN. This core technology is the foundation for “off-box” data protection schemes, including SnapVault and SnapMirror. SnapVault effectively propagates Snapshot copies to other NetApp storage devices—typically NearStore® systems—as a disk-to-disk backup solution for high-frequency “incrementals” that are accessed as point-in-time full backups. These, in turn, can be used for user-driven drag-and-drop file recovery and for periodic tape-based full backups without the pressures of production system “backup windows.” Open Systems SnapVault (OSSV) extends this functionality to third-party host-based file systems.

Asynchronous SnapMirror also draws from core Snapshot roots to provide data protection capabilities as part of a comprehensive disaster recovery/business continuance architecture.

RAID-DP

The core RAID technology has been augmented with a diagonal dual parity scheme, RAID-DP, that uses a second parity drive in a RAID group and “diagonal” parity computation to survive double-disk failures in a RAID group. The enhanced security of this RAID method allows for the creation of larger RAID groups (14+2 is common), effectively providing better than mirrored protection for user data with no additional

performance versus RAID 4. 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](#).

Increasing Database Agility for Test/Dev and QA

Replicating a large database for development, training, testing, or other purposes can be one of the most time-consuming tasks that a DBA can undertake. You have to carefully plan your methodology, provision enough storage to accommodate the copy, and then create a consistent replica of the data.

Learn how the use of NetApp technologies, including NetApp FlexClone and SnapMirror, simplifies the creation of local and remote database replicas. This can streamline the database application development, test, and deployment process to improve business agility.

Get the details. Read [Reducing Time-to-Deployment](#).

Advantages of NetApp FlexClone Technology in Database Environments

This five-minute swing bench demo shows:

- Performance benefits of running OLTP database loads on aggregated storage
- Ability to increase and decrease volumes in seconds
- Ability to create database clones in under a minute for testing

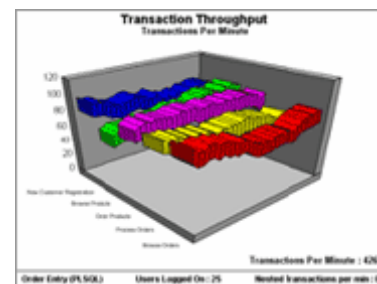
parity overhead (still 7:1) and negligible performance impact due to the multistripe write approach of WAFL.

FlexClone

FlexClone uses the same mechanism employed in Snapshot copies to create **writable** clones of volumes or LUNs. As with Snapshot copies, clones can be created nearly instantaneously with effectively no storage overhead; they share the same underlying storage blocks as their baseline volume/LUN. As the baseline and the clone diverge (for example, due to updates in the clone's data blocks), those new blocks and their blockmap pointers are written to disk, and the volume/LUN accumulates only the changed blocks.

This scheme affords space- and time-efficient, **writable** copies of file systems or LUNs that can be used for a range of purposes, including dev/test/QA; database reporting and analytics; and data warehouse extract, transform, and load.

Coupled with SnapMirror to create an incrementally propagated copy of data on a second storage system, FlexClone technology can be used to support these activities entirely "out-of-band" of the production storage system. This concept is depicted in figure 2.



[Launch the demo.](#)

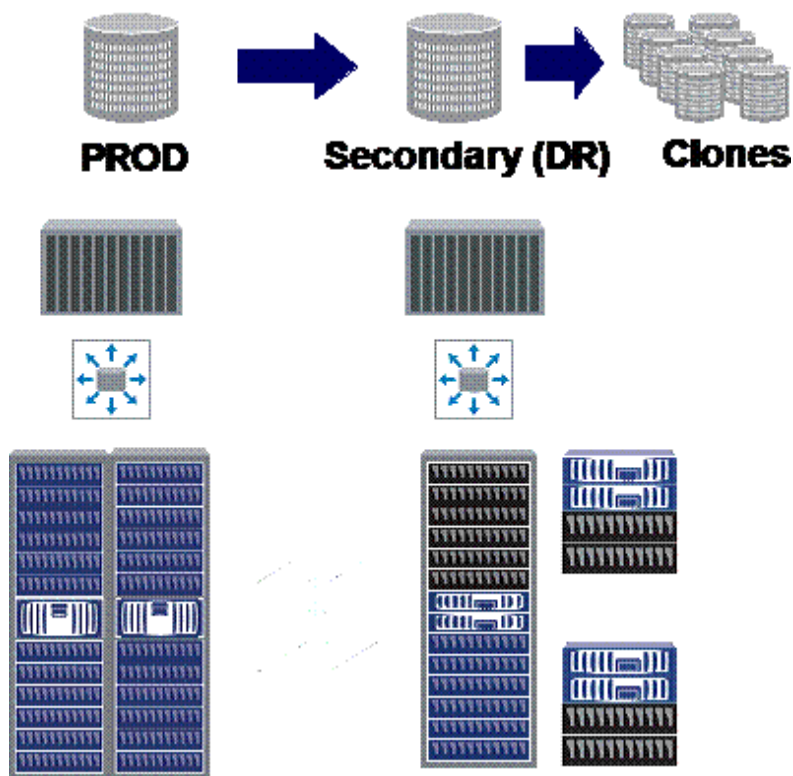


Figure 2) SnapMirror and FlexClone deployed in an out-of-band dev/test/QA scenario.

A-SIS

One of the newer features to evolve out of the NetApp genetic pool is Advanced Single Instance Storage, or A-SIS. Known also as block de-duplication, this feature uses the pointer and block management of WAFL to "squeeze out" duplicate blocks from the file system, replacing pointers to duplicate blocks with pointers to a common block. If files or volumes diverge after block de-duplication, the new (modified) blocks are stitched into the affected block map without impacting any other maps that may share the common block.

Conclusion

Many Network Appliance storage system features owe their existence to the core NetApp DNA: the unique combination of WAFL, RAID 4, NVRAM, and Snapshot that continues to fuel the evolution of the NetApp product line. And there are already additional extensions in the works that promise to deliver on the continued evolution this genetic blueprint affords.

[Comment on this article](#)