

Secrets to Shrinking Your Storage for VMware

By Michele Hope

July 2008

Sponsored By:



Copyright © The Storage Writer, Inc. 2008. All Rights Reserved.

The Storage Writer, Inc. + Independent Storage Insights, Technical Research and Reports
Phone: (770) 643-2050 + E-mail: mhope@thestoragewriter.com + Web: www.thestoragewriter.com

Table of Contents

Executive Summary	1
VMware: The Promise... ..	2
. . . and the Struggle to Achieve It	2
Understanding VMware's Duplicate Data from a Storage Perspective	3
NetApp Deduplication: A Hidden Asset for VMware Environments.....	5
Lessons in VMware Savings with NetApp Deduplication.....	8
NetApp Deduplication: How to Begin	11

Executive Summary

Virtual server environments powered by VMware have caused a revolution in the data center. Collapsing previously bloated and underutilized server and desktop infrastructures, VMware has led organizations to transform how they deploy and provision their applications, test and development copies, and virtual desktop environments that now benefit from this more fluid, dynamic architecture.

The promise of VMware has allowed many organizations to achieve big savings in server capital costs, including power, cooling and management. But, often, a different picture emerges on the storage side of this equation. As virtual machines (VMs) proliferate, storage also increases, often at an alarming rate for organizations seeking to shrink their overall IT infrastructure costs.

This brief looks at the challenge of harnessing storage growth within growing VMware environments, based on input from third parties, analysts, as well as current VMware users. It also explores reports from NetApp and NetApp customers that claim the ability to shrink and reclaim VMware customer storage footprints by anywhere from 50% to 90%.

How is this possible? Is it realistic to expect these kinds of savings in most customers' real-world environments?

The answer appears to lie in NetApp's deduplication functionality, now in use by a growing number of VMware installations. This brief explores potential deduplication savings and scenarios where this type of space capacity savings can best be reached.

VMware: The Promise...

Many IT organizations have begun to experience the benefits of consolidating their physical servers and applications onto a VMware-based virtual server infrastructure. Today, solutions like VMware's Virtual Infrastructure 3 (VI3) and its increasingly popular counterpart, Virtual Desktop Infrastructure (VDI), now make it possible to dramatically shrink the cost typically involved in acquiring, maintaining, powering and cooling growing numbers of physical servers or desktop PCs.

Instead, VMware gives customers the option to deploy many virtual servers *and* virtual desktops (also known as virtual machines, or VMs, that run as guest operating systems) -- all from within one underlying physical system. This underlying system is referred to as Virtual Servers within VI3 environments and Virtual Desktops in VDI environments. VMware users typically run about 6-10 VM guest OS's per physical server but this number often goes much higher. (NetApp sources report some customers running up to 70 VMs on a single server in VDI environments.)

As an organization's needs grow, and more virtual machines are provisioned, VMware gives customers the option to expand into a cluster of several ESX Servers which can quickly grow to support hundreds (or even thousands) of virtual machines.

From the server administrator's viewpoint, VMware's latest incarnations make it easier and faster than ever to provision more virtual machines with their own set of operating systems and applications already pre-installed. This is easily accomplished by first creating a set of pre-defined VM *templates* which act as golden images of the guest operating system and any associated applications. Such templates allow server administrators to quickly roll out, or "clone," new VMs to support the needs of emerging test, development or production applications, not to mention end users moving toward virtualized desktop applications. Another benefit, VMs can be more easily migrated, without downtime, from one physical server to another.

As organizations have begun to support more VMs and higher VMware-provided availability and data mobility features (such as Storage VMotion, DRS and HA), they made the necessary move to centralized storage and a network storage platform. Among other benefits, the use of network storage now allows multiple ESX Servers to access and use a common storage capacity pool, not to mention benefiting from other enterprise-class data availability and replication features found on the storage platform.

... and the Struggle to Achieve It

Here's where the challenge arises. While the use of VMware VMs causes the physical server footprint to shrink dramatically in the data center, the proliferation of VMs can simultaneously cause network storage capacity requirements to spiral upwards. In one case, a NetApp user reported his organization's roster of 1800 virtual servers and desktops had grown to require as much as 32TB of network storage capacity. NetApp deduplication functionality subsequently helped this user reduce his organization's virtual storage footprint by between 80 and 90%.

More VMs Equal More Duplicate Data

In a recent article on VMware, the CTO of GlassHouse Technologies, Jim Damoulakis, framed the storage issue as follows:

... Another important consideration is the increase of the volume of data created by virtual environments. The standard use of configuration templates for creating virtual machines results in numerous cookie-cutter VMs of a particular class that are nearly identical. Backing up and storing the myriad copies of nearly redundant data [also] consumes significant storage. . .¹

As Damoulakis pointed out, storage capacity issues are not just impacted on VMware's primary storage systems. Indeed, they also become a challenge when trying to achieve proper disk-based backup or replication of the growing virtual infrastructure -- especially where replication of very large data sets is not always feasible given the cost of bandwidth and WAN connections.

Dealing with Duplicate Data -- for Virtual Servers and Virtual Desktops

The issue of exponential storage growth for VMware environments is not just focused on those running virtual servers. In fact, the scale in which some companies deploy virtual desktops using VMware's Virtual Desktop Infrastructure, may bring storage capacity growth issues to a head earlier in the process. While virtual desktops may only be deployed with an initial 8-10 GB of storage capacity, companies that try to deploy hundreds (or thousands) of virtual desktops will soon see a storage impact.

In VMware's 2008 *Guide to Implementation Best Practices* for Virtual Desktop Infrastructure, Author Craig Cook of Long View Systems put it this way:

Given that each desktop typically will need about 8 GB of disk space, the storage infrastructure can quickly become cost prohibitive if you try to just extend your Virtual Infrastructure storage platform for use with VDI.²

Understanding VMware's Duplicate Data from a Storage Perspective

To set the groundwork for the rest of this brief, it's important to describe how duplicate data can proliferate so easily in VMware environments. This involves first defining a few concepts surrounding storage of virtual machine data.

What is a Virtual Machine?

According to VMware, a virtual machine is essentially a *"software container that bundles or 'encapsulates' a complete set of virtual hardware resources as well as an operating system and all its applications."*³ Virtual resources for each virtual machine include such items as virtual CPU, RAM, a Network Interface Card (NIC) and a virtual hard disk.

Which Files Represent a VM?

From a storage administrator's perspective, each virtual machine consists of multiple files stored in their own folder or subdirectory on the storage system. (For NetApp users running iSCSI or Fibre

Channel SAN storage protocols with VMware, the directory of files is stored on LUNs which typically reside on one or more NetApp flexible volumes. The contents of this directory are managed from within the VMware File System, or VMFS, datastore. In contrast, those running VMware under the NFS protocol can manage and view VM files and subdirectories directly from within the NFS-enabled NetApp flexible volume.)

Regardless of the protocol used to store the virtual machine, there remain two key files within each VM's "subdirectory." These are described in the following table.

Table 1. Key Files per Each VM.

File	Description
*.vmdk file	<p>The *.vmdk file represents all the contents of the virtual hard disk to the virtual machine. In fact, the virtual machine sees the *.vmdk file as just another SCSI-attached disk drive. One or more *.vmdk files may be in place for each virtual machine. Contents of a virtual machine's *.vmdk file(s) typically include:</p> <ul style="list-style-type: none"> ▪ Operating system data used by the virtual machine. This may include binary files, OS patches and drivers. This data is often referred to as the "guest OS". ▪ Applications running on the virtual machine, including application binary files, libraries, etc. These are often referred to as application images. ▪ Application data. ▪ Other temporary data, such as VM swap files, page files and user or system temp directories.
*.vmx file	The *.vmx file contains metadata and relevant details about each virtual machine, including its disk size, operating system in use and networking information.

(Other files are also found in each VM directory. These are less relevant to this brief, so will not be described here.)

How More VMs Create More Duplicate Data

Since the primary files for each VM are the *.vmdk files, VMware templates that "clone" each VM contain a lot of duplicate data. In fact, for VMs running the same operating system, each bit of OS data in the *.vmdk is identical from one VM to the next. Often, much of the application binaries from one VM's *.vmdk file to another VM (which is also running the same application) can also be identical. It's only when you begin looking at unique application data and other temporary data, do previously cloned VMs begin to look quite different from each other.

Figure 1 shows a conceptual view of the duplicate data you might find between one VM and another.

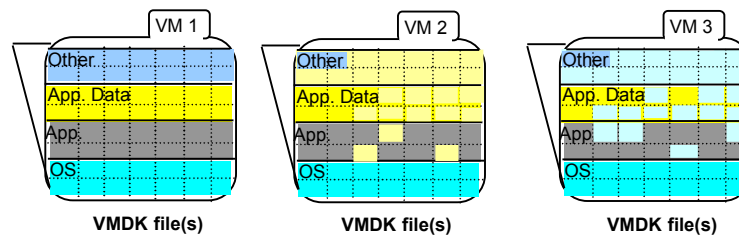


Figure 1. Example of Duplicate Data in each VM's VMDK files.

In a comparison between VM1, VM2 and VM3, each conceptual layer of data (OS, Application, etc.) is shown in a different color. Within each layer, different colored data “blocks” show data that's unique from one VM to the other. Each layer showing similar colored “blocks” between the VMs indicates duplicate data that would be considered a candidate for reduction via deduplication technology.

In the above example with just three VMs and no deduplication technology in use, one VM with a 20GB *.vmdk file would grow to three VMs, now needing 60GB. Hundreds or thousands of VMs simply compound the storage burden.

NetApp Deduplication: A Hidden Asset for VMware Environments

For some time, NetApp has reported significant success in reducing its VMware customers’ storage capacity requirements by as much as **50 to 90%**. The secret lies in the company’s deduplication feature which first became available with the release of NetApp’s Data ONTAP 7.2.2 storage systems. Originally known as A-SIS (for advanced single instance storage) deduplication, the company now refers to this functionality as simply *Deduplication for FAS*. This feature is widely available on NetApp’s current line of FAS systems.

Building on the Fundamentals of NetApp Snapshot and the WAFL Filesystem.

NetApp deduplication is an integral feature within Data ONTAP and the core NetApp WAFL filesystem. According to a recent NetApp article,⁴ NetApp deduplication functionality follows much the same process as existing NetApp Snapshot™ technology. They claim this is one reason deduplication accounts for relatively low system overhead during operation.

Specifically, NetApp deduplication builds on the concept of “multiple block referencing” developed years ago with NetApp snapshots. This architecture allows a single WAFL data block to represent multiple point-in-time copies. NetApp’s deduplication then adds the ability to identify unique vs. duplicate “digital fingerprints” associated with each 4KB block of data.

NetApp deduplication accesses and compares Data ONTAP metadata already containing the digital fingerprints associated with each 4K data block. After identifying duplicate fingerprints and performing a byte-by-byte scan to ensure the physical blocks are true duplicates, a simple block-redirect process moves the duplicate block to the “free pool”. It also references the duplicate block’s

data pointer to the original, previously stored block. The NetApp article refers to this process as “using one ‘physical’ data block to represent many ‘logical’ data blocks.”

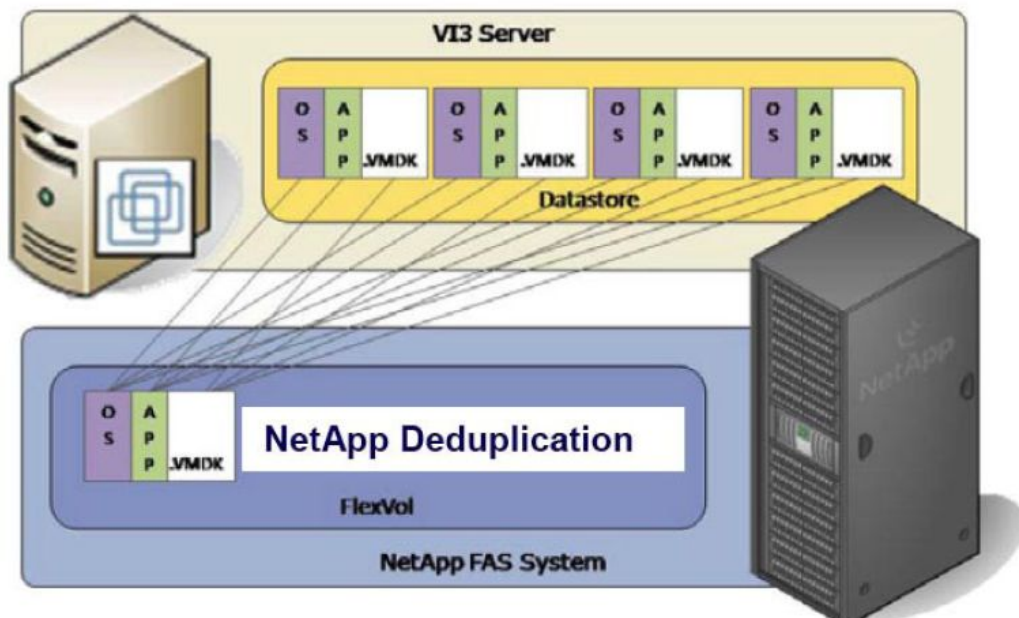
How Deduplication Works: An Overview

Since the digital fingerprints for each data block already exist in the WAFL filesystem, it’s an easy process to run NetApp deduplication. Instead of being focused on just backup data sets, NetApp deduplication runs on any data blocks contained within a deduplication-enabled NetApp flexible volume (also known as FlexVol™). In general, NetApp deduplication performs the following basic steps during operation:

1. Lists the digital fingerprints of block data residing in the deduplication-enabled volume.
2. Uses the fingerprints to compare and identify data blocks which are duplicates. According to NetApp, this keeps system overhead low during the process.
3. After identifying two or more duplicate blocks, it follows up with a byte-by-byte scan to confirm data in the block is an exact duplicate. (According to NetApp, this added scan removes the risk of “false positives” reported by other deduplication solutions while still minimizing the performance overhead required to perform deduplication.)
4. Moves any reference data pointers from the duplicate blocks to the remaining, “single instance” data block, while subsequently marking the duplicate blocks as “free.”

Operation with VMware

Some application environments lend themselves more readily to higher capacity savings and space reclamation with NetApp deduplication. In general, environments with high degrees of duplicate data like that found with VMware are often excellent candidates for deduplication, as shown in Figure 2.



Source: NetApp

Figure 2. NetApp Deduplication with VMware.

Some Important Points About Use with VMware

Unlike many deduplication solutions on the market, NetApp claims its deduplication functionality is unique in its ability to cover everything from deduplicating primary volumes to deduplication of secondary backups, archives and replicated data sets. The deduplication process also has a minimal performance impact while it runs, not requiring either the use of complex hash algorithms or look-up tables to preserve the data's "single-instance" status. NetApp deduplication features often touted include:

- The ability to deduplicate data on primary production storage volumes (currently a unique feature among vendors), as well as data being backed up or remotely replicated to secondary backup or archival volumes.
- The ability to deduplicate data residing on both new FlexVol volumes, as well as existing FlexVol volumes. This means flexible volumes with pre-existing applications and data, such as those commonly found in existing VMware environments, can still take advantage of deduplication.
- The ability to launch the deduplication process manually, on an automated (off-peak) schedule, or when the data capacity of a volume reaches a certain threshold. Deduplication can also be just as easily disabled.

VMware Caveats: While reports of high space savings with VMware are common among NetApp users, NetApp continues to caution that a VMware customer's specific deduplication "mileage" regarding VMware environment space savings and performance impact will vary. The OS and application portions of a *.vmdk file are likely to reclaim the most space from NetApp deduplication, especially for applications which are near replicas of each other. In contrast, a VM's unique application data and temporary files will not fare as well during the deduplication process. Certain applications fare much better than others.

General Recommendations: To maximize deduplication performance and space savings, NetApp recommends you:

- Store your VM's associated temporary files, such as VM swap files and page files, in another *.vmdk file on a separate NetApp FlexVol volume where NetApp deduplication is *not* enabled. This is separate from the deduplication-enabled volume containing your VM's primary OS and applications.
- If the VMware environment is supported by FC SAN or iSCSI SAN LUNs, pay close attention to how you configure LUNs on the NetApp flexible volume so as to maximize reclaimed space.⁵
- Become familiar with how NetApp deduplication works on volumes with existing or on-going NetApp Snapshot data.⁶
- Make sure your VMFS partitions have been properly aligned to the underlying storage.^{7, 8}

Efficiency Profile: *Report from the Blogosphere*

Christoph Biardzki is a NetApp customer who manages hundreds of terabytes of data at Germany's Leibniz Supercomputing Centre (LRZ), a mid-sized European HPC data center. He runs a blog entitled, "21st Century Storage."

One of his blog posts¹⁵ shared his impressions and experience using NetApp deduplication for VMware. Here's a summary of what he found:

Prior to Deduplication: One NetApp flexible volume was used to store a VMware cluster consisting of roughly 35 virtual machines (running both Microsoft Windows and Linux operating systems). With each VM needing from 5-25GB of storage, total storage used for all VMs was 504GB.

After Deduplication: The amount of storage capacity needed for the 35 VMs dropped to 130GB -- a 74% storage capacity savings.

Robert, a commenter to this post, also reported a space savings of 83% after using NetApp deduplication on roughly 300GB of VMs and templates associated with Microsoft Windows Server 2000, Windows XP and Microsoft Vista.

NetApp advises users to first assess the potential impact of running NetApp deduplication on each application used in their VMware environment. Applications requiring sequential reads or writes, for instance, do not make good candidates for deduplication. For details, see the *FAQ on NetApp deduplication for FAS*⁹ and a recent deduplication article¹⁰ from NetApp's *Tech OnTap Newsletter*.

Lessons in VMware Savings with NetApp Deduplication

Organizations who use NetApp deduplication with VMware tend to report significant storage capacity savings when performing key activities. These include:

- Cloning virtual machines
- Reclaiming storage space on existing storage volumes
- Replicating virtual machines to another location
- Backing up virtual machine data

The following subsections will detail some of the ways such capacity savings can be achieved.

Lesson #1: Cloning Virtual Machines

Most organizations use VMware's VirtualCenter to perform VM cloning from prior VM templates containing "golden images" of a pre-installed operating system and application.

After the VM cloning process is complete, a number of largely duplicate *.vmdk files are now present on the storage volume. Running NetApp deduplication on a flexible volume containing the *.vmdk files often yields high capacity savings. While NetApp estimates an average storage capacity savings of 70% using NetApp deduplication in VMware environments, the company reports having seen customer storage capacity savings range from 50% to 90%.

Cloning Virtual Desktops: In VMware's *VDI Guide to Implementation Best Practices*, Author Craig Cook lists deduplication as one of three storage options he recommends readers deploy to improve their ROI in growing Virtual Desktop Infrastructure environments. Displaying an example of NetApp FAS storage running NetApp deduplication, Cook notes the following:

Given that in a VDI environment you are deploying mirror image VMs from a standardized template, the vast majority of the information in the VMs [is] redundant. By eliminating all the redundant data through deduplication of the primary file systems, you can typically achieve a 50% or more reduction in the storage capacity required to support your virtual desktops, which significantly decreases the TCO and increases the ROI of the solution.¹¹

Efficiency Profile: *A Lesson in Cloning VMs*

One NetApp user at a financial institution has found another use for VMware with NetApp deduplication. When he needs his company's branch offices to test out a new business application, he no longer has to bring them all into the central office as he used to do. He also no longer needs to take up a lot of storage capacity to perform the testing.

With the help of VMware's VirtualCenter and NetApp features like deduplication and FlexClone, the user now quickly provisions the virtual machines he needs for each branch -- using very small amounts of storage. Here's how NetApp deduplication was able to help:

"Each environment required a 20GB file that we cloned six times into a 120GB NFS volume. Then we would deduplicate against that volume," he said, "Because the virtual machines are all identical copies, the deduplication would drop the total capacity back to 20GB."

His results with NetApp deduplication were so impressive he was able to deploy 35 virtual machines in just 50GB of storage capacity.

The user then extended his savings by deploying a total of more than 100 VMs. The results? "It still took only about 60GB," he said.

Lesson #2: Reclaiming Storage Space on Existing Storage Volumes

The process involved in reclaiming storage space “freed” by the NetApp deduplication process is fairly straightforward. In general, NetApp deduplication is designed to examine and discover all duplicate blocks in a NetApp FlexVol volume, then return any duplicates it finds back to the storage system. In effect, these duplicates are returned to the volume as “free storage” blocks for reuse. Any data pointers that previously pointed to the duplicate blocks will now point, instead, to the one remaining, identical block on the volume.

NetApp reports minimal system performance impact on volumes running NetApp deduplication as compared to volumes where deduplication is not enabled. While users typically try to run the deduplication process on weekends, evenings or off-peak hours, one user reported accidentally starting the deduplication process during prime production time and still reported little impact on their production systems.

Reclaiming Freed Blocks under NFS. VMware environments are supported by NetApp systems that run in either a file-based environment or a block-based LUN environment. After the deduplication process runs under NFS, freed data blocks are instantly recognized by the NetApp system and made available for use within the flexible volume.

Reclaiming Freed Blocks in a LUN Environment.

The process of viewing and reusing previously freed storage blocks in a LUN environment can require a few additional steps, especially if your LUNs have been configured with certain space guarantees. NetApp offers a report, *Configuring NetApp Deduplication with LUNs*, that describes how NetApp deduplication and space reclamation works under five different LUN configurations.¹²

Efficiency Profile: *Maximizing Storage Capacity*

Très Vance, manager of technical infrastructure at North Carolina-based Burt’s Bees,¹⁶ has seen the results of deduplication with VMware first-hand. After seeing the NetApp deduplication process reclaim at least 50% of the storage capacity allocated for new or current virtual machines, Très estimated he’d need three times the storage capacity from competing storage solutions compared to what he now has with NetApp. This is the amount of storage he’d likely need to support his company’s aggressive plans to move the rest of Burt’s Bees’ server applications to VMware. Instead, Très is confident his existing NetApp system with deduplication will be more than sufficient.

He explained the impact of deduplication as follows: “When I allocate a virtual machine in the VMware world, that means I can have 10 machines with the same operating system drive. NetApp allows me to basically create one copy of that operating system drive and nine other shortcuts or pointers.” he said.

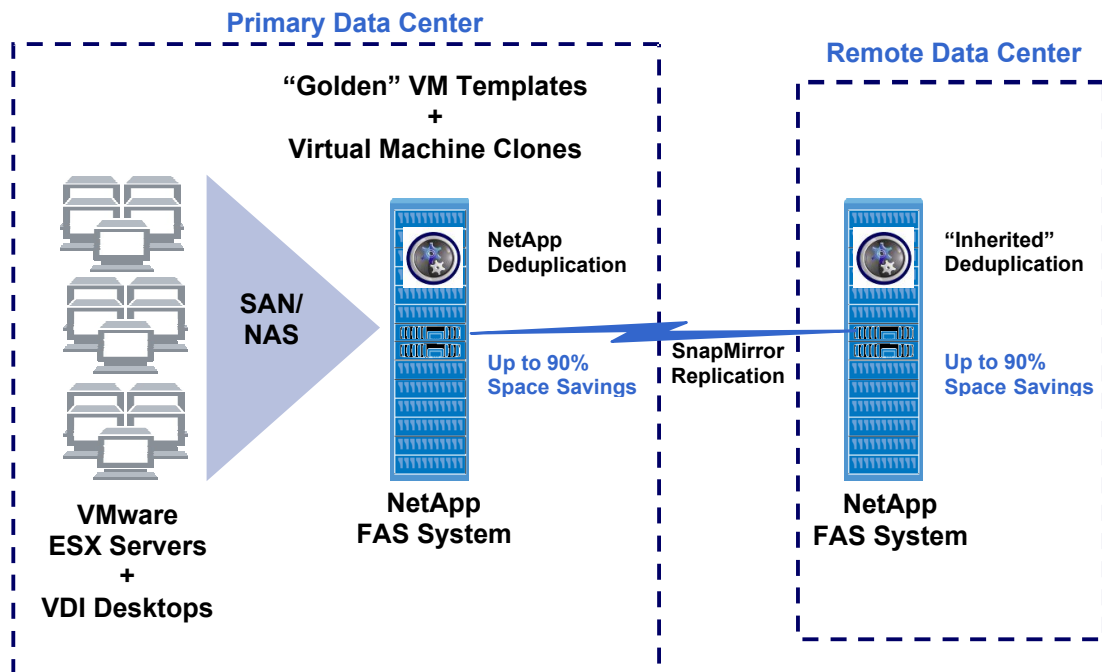
Très goes on to note the role he sees deduplication playing in growing data centers. “When you’re growing a business and adding new servers and new workflows, you need to know you’re not going to constantly be doing a one-to-one for all the storage. That can create a lot of heat in your data center you then have to displace with bigger air handlers. That whole environment snowballs pretty quickly.”

Lesson #3: Replicating Virtual Machines to Another Location

As more and more virtual machines begin to reside on a smaller footprint of physical servers, the need to replicate the underlying VM data to another location becomes more important as organizations seek to protect their most critical systems from a local or site-wide disaster.

Without deduplication, however, off-site replication may turn out to be a costly affair with largely duplicate data sets needing to be transmitted “across the wire”. In an article by NetApp Product and Partner Engineer Jeremy Merrill,¹³ Merrill describes the following deduplication scenario for environments using two NetApp systems -- one local and one remote. This scenario involves the use of NetApp SnapMirror[®] for remote replication between the sites. Figure 3 depicts this process which follows three basic steps:

1. The user deduplicates primary VMware volumes at the main data center, reclaiming up to 90% storage capacity.
2. Once VMware data is deduplicated, SnapMirror transmits the deduplicated 4K blocks of data (or subsequent changes to the original blocks) to the remote site. Each 4K block transmitted is unique. Duplicates are not transmitted, significantly shrinking the overall transmission costs.
3. Once the VMware replicated data resides at the remote site, you do not need to run deduplication again since all the benefits of deduplication are “inherited” from the primary system.



Source: NetApp

Figure 3. Deduplication Savings are “Inherited” by the Target Remote System During SnapMirror Replication.

Lesson #4: Backing Up Virtual Machines

Many NetApp customers have chosen to use NetApp SnapVault technology to perform efficient backups of the contents of one or more NetApp systems to another centralized NetApp storage system -- whether the centralized system resides somewhere on the local network or at a remote location. VMware environments also benefit from this type of backup technology which recognizes and sends only changed data blocks since the last scheduled backup.

Here's how NetApp deduplication adds greater efficiency to this type of backup scenario:

1. The user installs SnapVault and schedules SnapVault backup sessions according to his/her organization's requirements.
2. The user then enables deduplication on either the backup source, the SnapVault destination, or both.
3. Step back and watch how much storage capacity has been "freed" after deduplication.

How it Works: SnapVault is designed to be "deduplication-aware." For deduplication-enabled volumes, that means deduplication occurs automatically after each SnapVault session. Sites with deduplication running at both the source and destination will subsequently experience significant reclamation of storage capacity on both their primary storage as well as their backup storage system. Those running deduplication on their SnapVault system alone will see significant capacity savings with backup data.

NetApp Deduplication: How to Begin

Existing NetApp customers wanting to further explore use of NetApp deduplication in their environments just need to add two "no-charge" software licenses to their NetApp system: A NetApp NearStore license and a deduplication license. From there, deduplication is quickly enabled on any NetApp flexible volume via a simple CLI command. Other simple commands allow you to specify when and how often you want to run deduplication on the volume. Once deduplication is enabled on each volume, the volume's digital fingerprints will begin to be gathered and made ready for the deduplication process.

To learn more, go to:

<http://www.netapp.com/us/products/> and select the "Deduplication" link under the category of "Platform OS/Features". You can also refer to the End Note references at the end of this brief or to NetApp Technical Report 3505, "*NetApp Deduplication for FAS Deployment and Implementation Guide*."¹⁴

Efficiency Profile: *Replicating VMs*

In response to a blog posting by NetApp Global Systems Engineer Nick Triantos,¹⁷ one NetApp user indicated his organization had spent the past year using NetApp FAS3050 and 3070 systems to support over 800 VMs across 15 ESX Servers. That number has since grown to over 1000 VMs.

The user said his organization also uses SnapMirror to replicate VM data to a NetApp R200 system, with NetApp deduplication functionality enabled on the R200.

When asked by Nick about the space savings experienced on the R200 with deduplication, the user said, "We see well over 50% reduction of our VM volumes on the R200."

He also went on to note the following:

"We use the space savings to store 21+ days of snapshots which gives us the ability to restore full VMs from any day in the past 3 weeks... Just like tapes... But the restore (time in minutes) is just [an] NFS copy of the VM files..."

About This Report

This brief was commissioned by NetApp and involved the author conducting interviews with NetApp personnel and NetApp customers. Research was also performed on the part of the author via NetApp-internal sources and publicly available sources both within and outside of NetApp. Where possible, efforts were made by the author to attribute sources referenced throughout the brief and verify the accuracy of information presented herein. The author, Michele Hope, of the Storage Writer, Inc. is an independent storage industry commentator, writer and reporter with over 20 years of experience writing about IT topics. Michele's storage reports have appeared in venues ranging from InfoStor and Network World to FedTech Magazine and SearchStorage.com. She can be reached at mhope@thestoragewriter.com.

End Notes

- ¹ "Opinion: De-duplicating VMware," by Jim Damoulakis, Computerworld, March 12, 2008, <http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9068059>.
- ² "Virtual Desktop Infrastructure: A Guide to Implementation Best Practices," page 9, by VMware and Long View Systems, 2008, http://www.vmware.com/files/pdf/vdi_implementation_best_practices.pdf.
- ³ Source: VMware Web site, <http://www.vmware.com/overview/vmachine.html>.
- ⁴ "The Perfect Virtual Marriage: Deduplication and VMware," by Larry Freeman and Bill May, NetApp Tech OnTap newsletter, January 2008, <http://partners.netapp.com/go/techontap/mat/0108techontap.pdf>.
- ⁵ "Configuring NetApp Deduplication with LUNs," by Larry Freeman, Steve Botkin and Sajan Agrawal, NetApp, May 22, 2008, <http://communities.netapp.com/docs/DOC-1192>.
- ⁶ "Technical Report: NetApp Deduplication for FAS Deployment and Implementation Guide," Sections 4.4.3 and 4.4.4, by Bill May, April 16, 2008, NetApp Corporation, <http://media.netapp.com/documents/tr-3505.pdf>.
- ⁷ Ibid (refer to other sections in this NetApp Technical Report).
- ⁸ "Storage Block Alignment with VMware Virtual Infrastructure," by Tim Coste, NetApp Corp., July 2007, <http://media.netapp.com/documents/tr-3593.pdf>.
- ⁹ "NetApp Deduplication for FAS Frequently Asked Questions," NetApp Corp., April 9, 2008, <http://communities.netapp.com/docs/DOC-1060>.
- ¹⁰ "Deduplication Comes of Age, Version 2," by Blake Lewis, NetApp Tech OnTap Newsletter, July 2008, <http://www.netapp.com/us/communities/tech-ontap/dedupe-0708.html>.
- ¹¹ See Reference Note #2
- ¹² See Reference Note #5.
- ¹³ "Combining Deduplication and VMware DR," by Jeremy Merrill, NetApp Tech OnTap Newsletter, January 2008, http://partners.netapp.com/go/techontap/mat/dedup_VM_DR.html.
- ¹⁴ See Reference Note #6.
- ¹⁵ Source: "NetApp filesystem deduplication and VMware, Part 2," Blog posting by Christoph Biardzki, April 19, 2008, <http://21stcenturystorage.cebis.net/>.
- ¹⁶ Source: NetApp Success Story on Burt's Bees available at <http://www.netapp.com/us/library/customer-stories/>.
- ¹⁷ Source: Nick Triantos' blog posting appears at <http://storagefoo.blogspot.com/2007/09/vmware-over-nfs.html>.

NOTICE: The information, advice and/or recommendations made by the The Storage Writer, Inc. may include personal opinions of The Storage Writer, Inc. and others, which we believe to be accurate and reliable. Such information, advice and recommendations are made without warranty of any kind. The Storage Writer, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors which may appear in this document. All product names used and mentioned herein are the trademarks of their respective owners.