

HIGHLIGHTS

- ⌘ **Simplified SharePoint Backups**
- ⌘ **Streamlined DR for VMware ESX**
- ⌘ **Engineering View SnapManager/ SnapDrive**
- ⌘ **Resources:**
 - **Demo: SnapVault + Snapshot**
 - **Webcast: Veritas NetBackup**

NetApp A-SIS Deduplication: Top 10 Customer Questions

Get the answers to common customer questions regarding A-SIS deduplication licensing, configuration, performance, and support.

[More ⌘](#)



BLOGGING WITH DAVE

Dave Hitz, NetApp Founder and EVP

"It seems my recent blog post on whether iSCSI is SAN or NAS is offending people."

⌘ **Dave's Blog**

DRILL DOWN

- ⌘ **Webcast: Choosing a Disk-Based Solution for Veritas NetBackup™**
How IT teams are improving reliability, storage efficiency, backup, and recovery.
- ⌘ **Early Access: NetApp Snapshot™ and SnapVault® Technology Demo**
A conceptual overview.
- ⌘ **Lockheed Martin Reduces Risk, Speeds Exchange Upgrade**
Upgrading to Exchange Server 2007? Check out this Microsoft® case study.

TIPS FROM THE TRENCHES

Streamlined Disaster Recovery and More for VMware ESX Environments

Dave Cipparone, Customer Service Engineer, NetApp
David Edborg, Product and Partner Engineer, NetApp

Real-world examples of how IT teams have implemented cost-effective, heterogeneous DR and migrated from ESX 2.1 to ESX 3.0 with only 25 minutes of downtime. Plus, cloning to support test/dev from the DR site.

[More ⌘](#)

Three Ways NetApp Simplifies SharePoint® Backup and Restore

Jai Desai, Product and Partner Engineer, NetApp

Unique advantages of SnapManager for Microsoft Office SharePoint Server include:

- Fast backup (including metadata)
- Granular restore capabilities
- A SharePoint-centric interface

[More ⌘](#)



FEELING CREATIVE?

- ⌘ **"People who've never worked with NetApp would be surprised by [...]"**

Complete this sentence for a chance to win a NetApp Swiss Army watch and other prizes.



ENGINEERING TALK

Integrating Applications with Storage: Evolution of SnapDrive®/SnapManager®

Kostadis Roussos, Technical Director, NetApp Engineering

NetApp tools enable you to reconcile a host's notion of where blocks live with the storage system's notion, enabling the host to leverage NetApp backup, restore, and other capabilities.

In addition to standard tools for Microsoft Exchange Server, SQL Server™, SAP, Oracle, and SharePoint, the SnapDrive sdcli interface enables integration with any home-grown application. Includes a sample customer batch file.

[More ⌘](#)

FEEDBACK

NetApp A-SIS Deduplication: Top 10 Customer Questions Answered

In May, NetApp announced a new deduplication technology that can significantly increase the amount of data stored in a set amount of disk space: Advanced Single Instance Storage (A-SIS) deduplication. This technology is available (at no charge!) for NetApp NearStore® R200 and NearStore on FAS systems.

A recent NetApp TechTalk [Webcast](#) and [chat session](#) generated a substantial number of questions about the new technology, and we saw a very high response to last month's Tech OnTap [engineering perspective](#).

Below are answers to the top 10 customer questions about A-SIS deduplication:

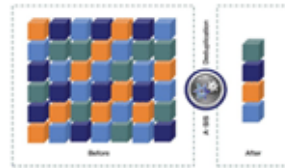


Figure 1. Before and after A-SIS Deduplication.

1. [What exactly is A-SIS deduplication?](#)
2. [How do I add the A-SIS deduplication capability to a system?](#)
3. [Why is an R200 or NearStore license required?](#)
4. [Are there any plans to remove the NearStore license requirement?](#)
5. [What do you mean by "light-duty" primary storage?](#)
6. [Can the system perform other operations while A-SIS deduplication is running?](#)
7. [Can I estimate my space savings before installing A-SIS deduplication?](#)
8. [Can previously written volume data be deduplicated?](#)
9. [Can I schedule the time when deduplication is run?](#)
10. [What kind of space savings can I expect?](#)

1. What exactly is A-SIS deduplication?

A-SIS deduplication is a general-purpose space reduction feature available on NearStore R200 systems and NearStore on FAS systems. When A-SIS deduplication is enabled, all data in the specified flexible volume can be scanned at intervals and duplicate blocks removed, resulting in reclaimed disk space. Note: A-SIS deduplication is not supported on R100, R150, FAS250, FAS270, or any products in the 800 or 900 families.

2. How do I add the A-SIS deduplication capability to a system?

The A-SIS license enables the deduplication capability. This license is available for R200 systems and for any FAS system that has a NearStore license installed. There is no charge for the A-SIS deduplication license. A-SIS deduplication requires a minimum Data ONTAP® version of 7.2.2 for FAS3000/6000 systems and 7.2.2L1 for FAS2000 systems.

3. Why is an R200 or NearStore license required to enable A-SIS deduplication?

Although A-SIS deduplication is application transparent, it has not been tested with mission-critical primary applications. By requiring a NearStore R200 system (or NearStore on FAS license) to enable A-SIS deduplication, we can help ensure that customers do not attempt to use A-SIS deduplication in performance-intensive application environments that have not been fully tested.

4. Are there any plans to remove the NearStore license requirement for A-SIS deduplication?

RELATED INFORMATION

- [Engineering Perspective: Deduplication Comes of Age](#)
- [Webcast: Cut Data Storage Use by As Much As 95%](#)
- [NetApp A-SIS Chat Transcripts](#)
- [Demo: NetApp A-SIS Technology](#)
- [Dave's Blog: How Deduplication Fits into Our Master Plan](#)

Engineering Perspective: Deduplication Comes of Age

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 accumulates 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.](#)

How Data Deduplication Fits into the NetApp Master Plan

Following is an excerpt from a recent post on NetApp founder Dave's blog:

Buying less storage is the small picture. The big picture is that we want to help customers create a disk-based copy for all of their primary storage. ...

Interesting things start to happen

**OnDemand
TechTalk Webcast:**

[Cut Data Storage Use by As Much As 95%: Using Deduplication to Drive Down Storage Costs](#)

We are continuously monitoring customer adoption of A-SIS deduplication in backup, archival, and light-duty primary storage environments. As we gain more experience with A-SIS deduplication in these environments, we fully expect its use to broaden. Until we gain that experience however, A-SIS deduplication will require the NearStore license.

5. What do you mean by "light-duty" primary storage? Is this appropriate for A-SIS deduplication?

What we mean by light-duty primary storage is volumes that contain primary (first copy) data, but that are not performance driven. Some examples of this would be user home directories, document directories, and application volumes that experience heavy I/O loads during the day but are quiescent at night and on weekends. These volumes might very well benefit from A-SIS deduplication if the system has the performance headroom to support the additional overhead imposed by A-SIS deduplication.

6. Can the system perform other operations while A-SIS deduplication is running?

Yes, A-SIS deduplication runs as a background process and the system can perform any other operation during this process.

7. Can I estimate my space savings before installing A-SIS deduplication?

Yes. A space estimation tool is available to NetApp SEs. This tool is a standalone application that operates on a Linux® client and will "crawl" through any NFS volume (either NetApp or non-NetApp NFS volumes up to 2TB) and will estimate the amount of space savings you will get with A-SIS deduplication.

8. Can previously written volume data be deduplicated?

Yes. A CLI command signals A-SIS deduplication to scan and deduplicate existing data on a volume. This command can be run at any time on a volume that contains previously written data, and we recommend it be run whenever A-SIS deduplication is first enabled on a volume.

9. Since A-SIS deduplication is run after the data is written to the volume, can I schedule the time when this deduplication is run?

Yes. A CLI command allows you to set individual A-SIS deduplication schedules for each volume.

10. What kind of space savings can I expect with A-SIS deduplication?

The space savings of any deduplication product is dependent on the number of duplicate objects that can be found and removed. A-SIS deduplication is no different. Based on internal testing and customer feedback, the chart below illustrates some sample space savings achieved by A-SIS deduplication in typical environments:

Dataset	Sample Space Savings Observed
Full System Backups*	20:1 (95%) over time
VMware Images	85%
Tech Pubs	50%
Software Archives	50%
Database	30 – 50%
Home Dirs	30 – 50%
Web & MS Office	30 – 45%
Oil & Gas Seismic Data	30%
E-mail Archive	20%

*In data backup environments, space savings grow over time as repetitive full backups are retained. For example, tests with CommVault Galaxy provided a 20:1 space reduction over time, assuming daily full backups with a 2% daily file modification rate.

when you create a disk-based copy of everything. Instead of doing searches on primary storage, which could hurt performance, why not search the secondary copy? If the people running decision support systems want their own copy of a critical database, why not clone the secondary instead of paying for a whole new copy? Why not create lots of cloned copies for the test and development team preparing to upgrade to the next version of Oracle or SAP?

When you create a copy of everything, and add functionality like Snapshot™ copies and clones, what you end up with is a smart copy infrastructure that can completely change the way you think about data management.

This won't happen overnight. We understand that. But anything that helps people reduce the cost of creating copies helps us achieve our vision more quickly. In the short run, data deduplication helps customers save space and save money, but what's more important is that by reducing the cost of copies, it helps us achieve our master plan.

[Read the full Dave's Blog post on deduplication.](#)



Dave Cipparone and David Edborg
NetApp

Combined, the authors of this article have over 25 years of data protection and business continuity experience. Both joined NetApp from Topio, Inc., and had previously worked together at Comdisco Continuity Services, where they helped many customers recover from disasters such as the World Trade Center collapse on 9/11. At NetApp, Dave Cipparone (left) is a data protection consulting systems engineer responsible for designing comprehensive data protection solutions, while David Edborg, a product and partner engineer, focuses almost exclusively on developing best practices and designing data protection and replication solutions for customers. Ran Pergamin, another NetApp product and partner engineer, also contributed to this article.

Streamlined Disaster Recovery and More for VMware ESX Environments

By Dave Cipparone and David Edborg

Deploying a virtual server infrastructure offers significant benefits, including a reduction in the number of physical servers and network ports needed as well as savings on floor space, maintenance contracts, and electricity. However, server consolidation based on VMware VI3 or other solutions may also create new management challenges and increases the criticality of backup/recovery and disaster recovery.

A recent Tech OnTap article titled [Five Ways to Use NetApp Snapshot™ Copies with VMware VI3](#) described how NetApp storage complements virtual server environments by streamlining backup copy creation, virtual machine recovery, and more. "That sounds great," commented several readers, "but what options are available when VMware VI3 is already deployed in combination with EMC, Hitachi, HP, or other storage platforms?"

In such scenarios, IT teams can continue to leverage their existing infrastructures while taking advantage of NetApp capabilities. One NetApp option is [ReplicatorX™ software](#). (See this month's sidebar for an alternative option using the NetApp V-Series product line.)

ReplicatorX (formerly Topio® Data Protection Suite [TDPS]) provides data replication and recovery across a broad range of locations, platforms, and storage systems and has specific advantages for customers who have or are deploying VMware ESX, including:

- Ability to use a wide combination of different storage systems and different connectivity methods (FC, iSCSI, DAS, internal/system) on the production side and at the DR site
- No disruption to production applications and minimal impact
- Block-based methodology that allows replication of VMware virtual machines, including system volumes (Windows® C:\ drive)
- Ability to nondisruptively leverage replicated data for DR testing, application test/dev, or other business purposes

The following examples describe how one customer implemented disaster recovery – plus plans to implement cloning for test and development – and how another customer streamlined a migration from ESX 2.1 to ESX 3.0 using this technology.

Cost-Effective DR for VMware VI3

NetApp recently worked with a large insurance company whose existing VMware environment used IBM DS4300 disk on a Fibre Channel SAN for primary storage. The team needed to reduce recovery time and recovery point objectives (RTOs/RPOs) for their disaster recovery solution. The solution had to include the ability to replicate not only the data volumes, but also the boot drives. The competition proposed a mirroring solution requiring channel extenders and high bandwidth. The bandwidth alone was

RELATED INFORMATION

- [Five Ways to Use NetApp Snapshot Copies with VMware VI3](#)
- [Love VMware, Hate Backups?](#)
- [ReplicatorX White Paper \(PDF\)](#)

A Full Suite of Data Protection Options

NetApp offers a comprehensive set of data protection solutions that includes high availability and clustered offerings, replication solutions, backup and recovery, disk-to-disk solutions, and virtual tape library solutions.

In addition to providing the required RPO/RTO, each solution enhances virtual and physical infrastructures by minimizing the complexity and cost associated with data protection.

Choosing a Replication Option

NetApp offers several tools for data replication: SnapMirror®, MetroCluster, and ReplicatorX.

If you only have NetApp storage systems, SnapMirror or MetroCluster is the best replication solution in most circumstances. Functionality ranges from asynchronous to fully synchronous and goes beyond disaster recovery with smart copies for DR testing, test/dev, data migration, data mining, offloaded backup, etc.

Learn more with:

- [NetApp SnapMirror Best Practices Guide. \(pdf\)](#)

cost prohibitive, and that did not include the budget to purchase (or staff to manage!) an identical storage system and Fibre Channel SAN switches and infrastructure at the DR site. The proposed solution would have locked the customer into that technology and would have increased the future cost, as both the production and DR site storage would need to be upgraded at the same time.

Instead, the IT team chose to replicate VMware volumes using ReplicatorX and implement an iSCSI-based NetApp SAN using a FAS3020 storage system with FlexClone® at the disaster recovery facility. The ability to replicate from the Fibre Channel environment at the primary site to iSCSI at the DR site (instead of Fibre Channel) offered significant costs savings and reduced complexity for the DR site.

The team uses ReplicatorX to replicate all VMware guest volumes, including C:\ drives, from IBM storage at the primary location to the NetApp system at the DR site. ReplicatorX clients on the production side transfer data over an IP network to a replication server at the DR site. All the "heavy lifting" of replication is done by the replication server at the remote site, rather than the replication clients at the primary site, minimizing the impact to production servers.

With ReplicatorX you can size bandwidth to meet a desired RPO. Bandwidth sized to write peaks can achieve near synchronous replication. When adaptive data reduction techniques are used, bandwidth cost can be traded off against a slightly higher RPO. A compression option can further reduce bandwidth requirements.

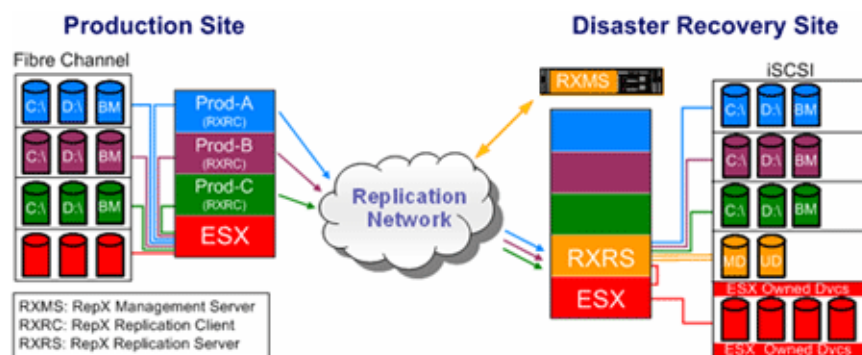


Figure 1) Typical VMware V2V DR scenario with ReplicatorX. All of the DR iSCSI drives are mapped to the RXRS from the storage system, bypassing ESX. The recovery servers are defined but powered off.

For this IT team, the capital and operating expense cost savings were substantial with the NetApp, iSCSI, and VMware solution, but the real value of the solution for the team is the ability to nondisruptively clone data to test the DR environment. The combination of ReplicatorX, NetApp FlexClone technology, and VMware VI3 made this possible.

DR best practices mandate that DR testing should not be performed on replication target volumes in case a real disaster occurs during testing.

ReplicatorX is closely integrated with the NetApp FlexClone technology available on all NetApp storage systems, which can create a virtual clone of a storage volume that only consumes additional disk space as changes are made. When changes are made to either the clone or the original, new data is written to a new location and simply mapped to the current image of the virtual volumes.

This is a unique NetApp feature with distinct advantages over other "copy-on-write" snapshot technologies.

To perform a DR test, an administrator can simply freeze the target volumes without sacrificing RPO, create FlexClone volumes of all VMware volumes, unfreeze replication, and then bring up the virtual machines using the FlexClone volumes while replication continues undisturbed. This gives customers greater confidence that they will be able to recover from a disaster should it be necessary.

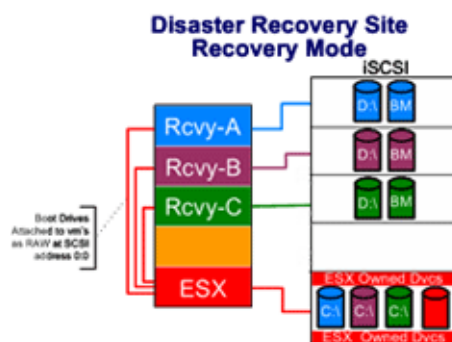


Figure 2) Figure 1 after recovery. Boot drives are mapped through ESX, and data drives are mapped directly from the storage system to the recovery server's iSCSI initiator, bypassing ESX; this simplifies the recovery process while increasing VMware guest deployment flexibility.

- [NetApp SnapMirror: The Director's Cut](#) (article plus technical Q&A).

Choose ReplicatorX if you have third-party primary storage and you want a tiered storage solution (different, less expensive storage on the back end) or you have a variety of primary storage systems and need a solution that can cover all of them.

Also, you should choose ReplicatorX if you have federated applications (SAP, DB2 IEEE, etc.) where data needs to be recovered at the exact same time across multiple servers. The ReplicatorX solution has patented global clock synchronization technology that guarantees write order consistency across multiple servers at the target location.

Read the white paper [Data Replication and Recovery over Any Distance with ReplicatorX](#). (pdf)

Using V-Series with Third-Party Storage

An alternative option to ReplicatorX in environments with third-party primary storage is the NetApp V-series virtualization system.

Because ReplicatorX is a software-only solution, it may be easier to deploy into established environments. V-Series is a SAN virtualization engine that front ends existing storage arrays so that you can leverage the full suite of NetApp storage features in your production environment – including SnapMirror for replication.

Several recent Tech OnTap articles highlighted use of V-Series:

- [V-Series Deployments in a Variety of Storage Environments](#).
- [V-Series to Support a Nonvirtual SQL Server Environment](#).

Love VMware but Hate Backups?

While no one disputes the obvious benefits of VMware and server virtualization, the radical change in architecture has significant implications for backup, HA, disaster recovery, and so on.

NetApp VMware expert Vaughn Stewart tackles some of the thorniest questions posed by end users. Examples include:

- The fact that a single service disruption could affect my entire environment is keeping me awake at night. Do I have to live with this level of insecurity?

In a real disaster, as a best practice NetApp recommends that you follow the same process and create FlexClone volumes of replicated volumes prior to starting your virtual machines just in case anything should go wrong during the recovery.

Using ReplicatorX/FlexClone to Support Test/Dev

The same IT team is also considering taking advantage of the ability to create clones of production data at the DR site as an inexpensive way to do test and development work without impacting production operations. All virtual machines can be started using FlexClone volumes of the target volumes without disrupting ongoing replication or the production VMware environment. If a failure occurs, all active virtual machines at the DR site can be rapidly switched to use the latest replicated volumes and brought up in production mode.



Figure 3) DR scenario in which replicated FlexClone volumes are used for test/dev during normal operation.

Nondisruptive VMware ESX 2.1 to ESX 3.0 Migration

In another case, an American financial services company needed to upgrade active VMs running Windows applications from ESX 2.1 to ESX 3.0. Using a manual data copy procedure would have required 20 hours of downtime on critical systems, which was far more than the customer's ongoing business operations could tolerate.

During an internal brainstorming process, the IT team realized that it would be possible to complete the upgrade with minimal downtime using ReplicatorX, which was already in use for disaster recovery. Taking advantage of the one-to-many replication capability (one volume to many destinations) of ReplicatorX, it is possible to create a replica of the production data in the new ESX environment, while continuing to replicate data to the DR site.

The IT team contacted NetApp support to ensure they understood all the details and to develop a formal process for migrating Windows servers in a VMware environment. Once the necessary volumes have been replicated, cutover from an existing VMware environment to a new VMware environment is a straightforward process:

1. Shut down the application(s) in the existing (production) environment.
2. Use the ReplicatorX GUI to verify that all of the production data has been replicated to the new environment, that write activity on the source volumes has been quiesced, and that no transactions are in flight on the network.
3. Freeze the replicated data in the new VMware environment.
4. Shut down the old production servers.
5. Prepare volumes for mounting.
6. Shut down the replication server.
7. Remap the virtual disks from the replication server to the newly defined ESX servers.
8. Start the ESX servers and virtual machines. Since the system disks have been replicated, the new servers automatically come up with the identity of the old servers, including software, processes, and schedules.

- My backup processes haven't changed, but suddenly I'm spending almost all my time managing backups. How do I stop the madness?

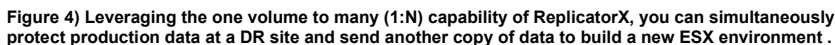
[Read Love VMware, Hate Backups?](#)

ReplicatorX Impact on Host Performance

One of the biggest concerns with host-based replication software is that the additional burden the software places on the host may steal too many CPU cycles and impact application performance.

To assess the impact of ReplicatorX running at host level, tests were executed using a variety of I/O workloads with and without the ReplicatorX client software. These tests show that even under the most stringent I/O, ReplicatorX increased CPU utilization by only 3%. For more typical I/O workloads, the ReplicatorX client software had negligible CPU impact.

[View the complete report.](#) (pdf)



For this customer, the ability to minimize downtime for critical virtual machines was the key factor. Because of the ReplicatorX architecture, there was no disruption to ongoing work while the data was being replicated in preparation for cutting over, and it was possible to replicate everything, including system disks, so configuration time was minimized.

Whatever your replication needs, Network Appliance offers solutions that can address them with minimal disruption to production environments. ReplicatorX also works in other virtual and physical environments to support almost any application requirement using any storage as the source and target. Although this article highlights the use of NetApp storage in the DR site, it is not required.



Jai Desai

Product and Partner Engineer, NetApp

Jai spent more than four years working in all areas of QA for Cisco Systems prior to joining NetApp as a requirements engineer in April 2005. In this role, he documented deployments at top NetApp accounts and worked with NetApp engineering to simulate these environments (and planned architectural modifications) during new release testing. In his current role, Jai focuses on the Data Suite in the NetApp manageability software family. He writes NetApp best practice guides for products such as Protection Manager and SnapManager® for Microsoft® Office SharePoint® Server and consults with NetApp SE teams around the globe regarding customer implementations.

Three Ways NetApp Simplifies SharePoint Backup and Restore

By Jai Desai

Microsoft Office SharePoint Server 2007 includes content management, information search, portal services, and other capabilities in a single software package. SharePoint Server 2007 combines the features of SharePoint Portal Server 2003 and Content Management Server with new capabilities, extending the functionality of the Microsoft Office suite to centralize storage of documents and other content for simplified management and improved Web-based collaboration

SharePoint Server 2007 offers many advantages over SharePoint 2003, and usage is accelerating rapidly. IT teams are deploying it to centralize important content from remote locations and to get important documents off of individual desktops and laptops for improved accessibility as well as better data protection and security. SharePoint Server 2007 search capabilities make it easier to locate data, including specific versions of stored files. In fact, Microsoft is actively encouraging customers to move away from Exchange Public Folders for sharing Office documents and other files.

Naturally, the more corporate documents you pull together in a single location, the more business critical the repository becomes. As customers use SharePoint for business-critical collaboration and content management, storage availability, accessibility, and performance are becoming critical, as is data protection – a particular pain point for SharePoint Server 2007.

With the newly released [SnapManager for Microsoft Office SharePoint Server](#) (SnapManager), NetApp helps streamline data management in SharePoint Server 2007 environments through:

- Fast backup of the SharePoint content database, including metadata
- Granular recovery of content, from a single item to the entire content database
- A SharePoint-centric user interface that simplifies data management tasks for the SharePoint administrator

These points are explored in detail in the following sections.

Benefit #1: Fast, Space-Efficient Backup

With almost any application, two things are critical when it comes to backup:

- Speed – you can't afford to have your SharePoint application offline for extended periods during backups
- Space efficiency – the less space your backup takes, the more backups you can retain

Unlike traditional backups, NetApp snapshot-based backup sets won't consume additional disk space unless changes are made to the content being protected. For example, suppose that an existing file is modified and saved to disk. A traditional backup would copy the entire file into the backup set. A NetApp Snapshot™ backup

RELATED INFORMATION

- [SnapManager for SharePoint](#)
- [Engineering Perspective: Evolution of SnapManager/SnapDrive](#)
- [Streamlined Backup/DR Admin](#)
- [Avoid Death by Scale](#)

Make the Switch to Proactive Data Management

Most storage administrators would prefer to spend their time on value-added tasks that directly affect business operations, but instead are spending an increasing portion of every day responding to one-off requests from system administrators, DBAs, and application administrators.

NetApp simplifies the data management problem by enabling systems administrators and DBAs to manage their own data and self-serve basic requests. The storage administrator retains control of the process through the ability to govern the operations that other data owners are allowed to perform.

In a recent Q&A, NetApp general manager Dave Kresse explained [the NetApp philosophy of Integrated Data Management](#).

Simplified Backup and Replication Management with NetApp Protection Manager

Protection Manager is an intuitive, policy-based management application for NetApp disk-based backup and replication technologies, including SnapMirror, SnapVault, and Open Systems SnapVault.

will not copy the entire file, but will only keep file system pointers to the blocks underlying the file. The old changed blocks are not deleted, but remain on disk, protected until the backup set that is dependent on them has been deleted. Because unchanged blocks are not recopied to disk for backup, the NetApp approach to backing up data is very space efficient and can enable 250+ Snapshot copies in a single storage volume.

Unless you know the mapping between files and blocks, however, you can't recover a specific SharePoint file or version of the file from a backup set. (This is true whether you're using NetApp storage or another vendor's array.) This is where SnapManager comes into play. SnapManager indexes SharePoint content and maps it to the underlying blocks on the NetApp storage system. This enables fast and highly efficient backups without sacrificing recovery flexibility, which I'll explore in detail in the next section.

To create a Snapshot copy, SnapManager first ensures that the SharePoint repository resides on NetApp storage and that all cached data and metadata has been flushed from memory to disk (a consistency check point). Then SnapManager initiates a Snapshot copy to rapidly back up the underlying SQL Server™ data as well as associated metadata. The Snapshot process takes only a few seconds, so normal operations should not be disrupted. (By comparison, the native backup capability in SharePoint Server 2007 is very resource intensive, and Microsoft doesn't recommend running it while a system is handling transactions.)

Once a Snapshot copy is created, you have multiple options.

- You can retain the Snapshot copy on primary storage.
- You can back up the Snapshot copy to tape.
- NetApp SnapVault® can be used to back up the Snapshot copy to secondary disk storage.
- NetApp SnapMirror® can be used to replicate the Snapshot copy to a disaster recovery site for end-to-end data protection.

All of these options occur on the storage system with no impact to the SharePoint application. By moving the protection of the SharePoint data off the application server, NetApp enables less risk of unforeseen downtime.

As an example of the benefit of fast and nondisruptive backup, suppose that a user accidentally infects the SharePoint document repository with a virus – an unfortunately common event in Microsoft Office environments. Recovery from a traditional nightly backup might result in the loss of up to a day's worth of changes and could take hours. Having hourly backup Snapshot copies enables you to revert to the backup set created immediately before the infection occurred or to simply recover an uninfected version of the affected file or files (depending on the severity of the problem), all in minutes.

Benefit #2: Granular Recovery

The native SharePoint Server 2007 recovery capability is limited to restoration of Content Database information at the SharePoint site level. To recover a single file, you have to restore the Content Database to a nonproduction server and then manually copy the file back to the production server, a time-consuming process that results in loss of the SharePoint metadata.

Using SnapManager enables SharePoint Server 2007 recovery at any level of granularity, from the database level all the way down to a single item, nondisruptively, while the SharePoint application server is online. For instance, suppose that you lost a particular version of a file from several months ago. With SnapManager, you can search for the file by name. When you search on the file name, you see all the Snapshot copies that contain the file so that you can choose the appropriate version of the file and recover it. File metadata – which can include date/time created, date/time modified (all versions that have been checked in), and security or access rights – is recovered along with the file. The entire process is nondisruptive to other SharePoint activity.

Similarly, if you are restoring a particular folder or subsite, you can navigate to it in a SharePoint tree structure in the user interface and select the desired object to be recovered. Because you are restoring the data from a Snapshot copy on disk, the recovery proceeds much faster than with other alternatives.

If you are mirroring to a remote site for disaster recovery, breaking the SnapMirror relationship makes the destination writeable. User access can be shifted to SharePoint volumes at the remote site until activity is restored at the primary site. This minimizes user disruption and data loss.

This tool enables administrators to apply predefined policies to their data, thereby eliminating ambiguity and the potential for error inherent in manual management.

A new demo shows how Protection Manager solves three common management challenges:

- How do you make sure that everything is protected when data is distributed everywhere?
- How do you scale your data protection environment without spending all day on tedious manual tasks?
- How do you rapidly roll out global changes across all sites and systems?

[Watch the demo.](#)

[Read Three Backup and Replication Management Challenges Solved](#)

SnapManager for Oracle

For Oracle DBAs, backup and recovery is a complex task that may require frequent negotiations with both system and storage administrators. SnapManager for Oracle allows DBAs to manage their own data while offloading data management work from Oracle® servers to NetApp storage.

From a central management system, a DBA can monitor the state of all Oracle backups across multiple databases, even if they are running on different servers and operating systems.

In a recent article, NetApp database architect Alvin Richards explored the close integration between SnapManager for Oracle and Oracle, including integration with critical Oracle utilities such as RMON and ASM.

[Read the article.](#)

Benefit #3: Single, SharePoint-Centric Interface

Another benefit of SnapManager is that it has a centralized interface that is very friendly and familiar to the SharePoint administrator; you don't have to know anything about how the Microsoft SQL Server database that underlies SharePoint Server 2007 is laid out, or anything about the underlying NetApp storage system. The look and feel of the GUI is the same as if you were browsing your SharePoint environment. You simply locate and select the items you want to back up or recover. SnapManager translates those actions to appropriate underlying actions on the database server and storage system.

In addition to backup and recover, SnapManager also provides the ability to create backup and retention policies as well as create and manage SharePoint users, groups, and permissions.

You can create up to six backup schedules and associate a retention policy with each schedule. For instance, you might have a separate schedule with a separate retention policy for hourly, daily, weekly, and monthly Snapshot copies. A verify option instructs SnapManager to verify any backup after it is created. You also have an indexing capability during backup.

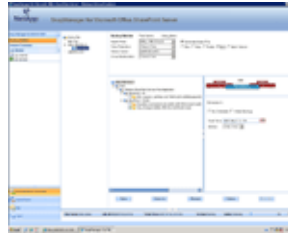


Figure 1) SnapManager for Microsoft Office SharePoint Server interface.

From the centralized SnapManager GUI, you can also manage a SharePoint infrastructure spanning multiple sites. A plan can be defined for each site that includes site-specific backup and retention policies with six unique backup schedules per site, as described earlier. A built-in logging mechanism enables you to run reports on the success of all backup and recover jobs.

Robust Storage Solutions for SharePoint Server 2007

There's one last thing I should point out. Although SnapManager for Microsoft Office SharePoint Server is new, it leverages the underlying NetApp SnapManager approach to application management (see the related article Integrating Applications with Storage: The Evolution of SnapDrive® and SnapManager) and specific capabilities associated with SnapManager for SQL Server, which has been deployed in hundreds of production environments.

With an easy-to-use and intuitive interface and the ability to back up or recover any object in your SharePoint Server 2007 environment without losing important file metadata, SnapManager helps take the complexity (and stress!) out of SharePoint data management. Used in conjunction with NetApp resiliency and data protection features such as RAID-DP™, SnapMirror, SnapVault, and optional encryption, SnapManager can help you protect SharePoint data against almost any threat.



Kostadis Roussos

Technical Director, Storage Management and Application Integration, NetApp Engineering

Kostadis joined NetApp in 1999 from SGI. For the next five years he worked on streaming media caching as part of the NetCache® organization and then file caching as part of DNFS and later PFS. In 2004 he discovered manageability. As an individual contributor in the manageability group, he worked on several releases of DataFabric® Manager. As a technical director, he spends most of his time tracking the various technology, customer, and competitive trends in applications and operating systems that affect the simplicity of data management. He uses that understanding to help the Storage Management and Application Integration Business Unit define its roadmap.

Integrating Applications with Storage: Evolution of SnapDrive® & SnapManager®

By Kostadis Roussos

The IT challenge is how to provision functional services such as e-mail with the right data service attached at the right cost. Or to put it differently, e-mail has to have the right number of backups, the CEO's e-mail must not be deleted, and the e-mail service has to tolerate the right number of failures without breaking the budget. Failure to do so results in down time that immediately affects business, especially for mission-critical applications like Microsoft® Exchange Server.

Data ONTAP® provides many of the basic building blocks that can be used to deliver the appropriate data service at a fraction of the cost. For example, consider an application that requires hourly backups. Using Snapshot™, a storage administrator can perform a near instantaneous backup with minimal downtime.

The problem is that for most applications, any data management operation such as backup must be coordinated with the application and in many cases with the underlying operating system. Without such coordination, the backup is only crash consistent.

To address this challenge, NetApp has created a suite of intelligent data management offerings that allow applications and storage to coordinate activities to make backup fast and space efficient, speed the restore process, and simplify common data management tasks. NetApp users today can choose from a suite of five SnapManager offerings that coordinate data management activities with many of the most popular enterprise software programs.

A Short History: Why Integrate Storage with Microsoft Exchange?

Our application integration strategy and SnapManager for Exchange product began life when NetApp proactively partnered with Microsoft on new ways that we could add value to Microsoft applications. The Microsoft Exchange team, in particular, learned about our Snapshot technology and asked whether we could develop a script that would allow Exchange to be backed up very frequently using Snapshot copies. Dennis Chapman, a senior technical director at NetApp, was asked to look into the problem. Dennis quickly determined that a scripted solution was possible, but he observed that a scripted solution would force Exchange to shut down while the Snapshot copy was in progress. In his exploration he discovered that it was possible to make Exchange think that taking a Snapshot copy was just a really, really fast tape backup. The advantage of this approach was that Exchange could stay online during the entire backup process.

Based on this insight, NetApp built the original SnapManager for Exchange as a tool targeted to Exchange administrators that enabled very fast backups of Exchange using Snapshot copies. Over the next several years, as Exchange and Data ONTAP evolved, so did SnapManager for Exchange.

The Creation of SnapManager and SnapDrive

RELATED INFORMATION

- ✦ [Three Ways NetApp Simplifies SharePoint Backup and Restore](#)
- ✦ [Empowering DBAs with SnapManager for Oracle](#)
- ✦ [Streamlining Backup and Disaster Recovery Management](#)
- ✦ [SnapDrive Technical Overview \(PDF\)](#)

Make the Switch to Proactive Data Management

Most storage administrators would prefer to spend their time on value-added tasks that directly affect business operations, but instead are spending an increasing portion of every day responding to one-off requests from system administrators, DBAs, and application administrators.

NetApp simplifies the data management problem by enabling systems administrators and DBAs to manage their own data and self-serve basic requests. The storage administrator retains control of the process through the ability to govern the operations that other data owners are allowed to perform.

In a recent Q&A, NetApp general manager Dave Kresse explained [the NetApp philosophy of Integrated Data Management](#).

SharePoint Data Management Simplified

Data protection in SharePoint Server 2007 can be complex. Using SnapManager for Microsoft Office SharePoint Server, you can create detailed backup schedules that use NetApp Snapshot to ensure that

The big change that occurred for SnapManager for Exchange was the introduction of FC and iSCSI support in Data ONTAP 6.3. Until that point, SnapManager for Exchange was tightly coupled with Microsoft Windows®, Exchange, and Data ONTAP. One engineering challenge was that every release of Windows, Exchange, or Data ONTAP required a new release of SnapManager for Exchange. The other challenge was that the team wanted to add more instances of SnapManager, in particular SnapManager for SQL Server™. The team decided at that point to split SnapManager for Exchange into two parts:

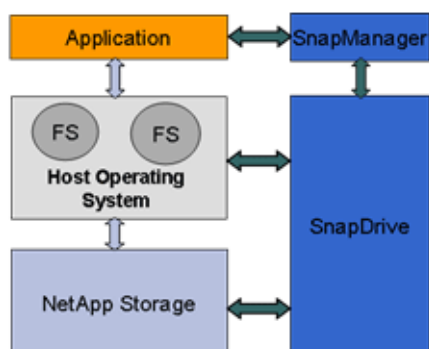


Figure 1) Roles of SnapDrive and SnapManager.

- SnapDrive understands and coordinates activities between the host operating system and file system and NetApp storage.
- SnapManager understands and coordinates with the application (and with SnapDrive).

Splitting the functionality makes it possible to leverage the underlying capabilities of SnapDrive in the construction of new SnapManager offerings tailored to specific applications, thereby simplifying the development process. SnapDrive and SnapManager are able to reconcile the host's notion of where blocks live with the storage system's notion so the host can leverage all the underlying capabilities of NetApp storage such as Snapshot, FlexVol®, FlexClone®, SnapMirror®, etc.

Host/Storage Integration – Automatic, Near-Instant Operations *NetApp Tool: SnapDrive*

Applications running on a host with SnapDrive installed (even without an appropriate SnapManager product) are able to use LUNs on NetApp storage systems as if they were local resources, while benefiting from underlying NetApp technologies such as Snapshot and cloning. Most tasks occur in seconds, so operations are not disrupted, and the operator doesn't have to know anything about NetApp storage or the specific networking technology (such as iSCSI or Fibre Channel).

The current generation of SnapDrive products includes SnapDrive for Windows and SnapDrive for UNIX® (with support for Solaris™, HP-UX, AIX, Red Hat, and SUSE Linux®). Both solutions support iSCSI and Fibre Channel transports with multipathing for performance and reliability. SnapDrive for UNIX also works over NFS.

SnapDrive helps host administrators provision storage and manage it directly from the host, enabling them to flexibly define backup policies and resize storage on the fly without any disruption of application service. SnapDrive understands the operating system, volume manager, and file system details necessary to coordinate activities with the host and to appropriately map host file systems to NetApp storage and WAFL®. Both versions simplify storage and data management by hiding the complexity of steps that must be executed on both the storage system and the host system, and by removing the dependency on the storage administrator.

Key SnapDrive functionality includes storage provisioning on the host, consistent Snapshot copies, and rapid data recovery from Snapshot copies. SnapDrive complements the native file system and volume manager technology and integrates seamlessly with clustering technology for high availability. Role-based access control (RBAC) prevents administrators from performing actions on LUNs for which they don't have permission.

Making these operations transparent to the user is not as simple as you might think. For instance, to create a consistent Snapshot copy, you must ensure that all file system data and metadata is flushed to disk and the file system is quiescent. That

critical content and metadata are always protected.

Granular recovery capabilities allow you to quickly recover a file, a folder, or an entire site, without taking SharePoint offline. A SharePoint-centric interface hides the underlying complexities to make these operations and more a snap.

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

SnapManager for Oracle

For Oracle DBAs, backup and recovery is a complex task that may require frequent negotiations with both system and storage administrators. SnapManager for Oracle allows DBAs to manage their own data while offloading data management work from Oracle servers to NetApp storage.

From a central management system, a DBA can monitor the state of all Oracle backups across multiple databases, even if they are running on different servers and operating systems.

In a recent article, NetApp database architect Alvin Richards explored the close integration between SnapManager for Oracle and Oracle, including integration with critical Oracle utilities such as RMAN and ASM.

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

Information-Optimized Backups

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

Over the past five years, however, disk capacities have grown 20x, with a scant fourfold increase in disk speed. This has made it more time consuming to move entire copies of data off of disk during a full backup. Today, many companies are turning to smarter backup solutions that:

- Shrink backup windows with Snapshot backups and replicate only incremental data changes
- Reduce backup storage consumption by eliminating redundancy in backup files
- Ensure reliable, recoverable backups by validating backups as they are created

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

requires the following general steps:

1. Create a list of LUNs to be included in the Snapshot copy. A Snapshot copy can span volumes or even multiple storage systems.
2. Stop activity on all affected volumes.
3. Flush data from the host operating system for all volumes in the Snapshot copy. (In Windows 2003, this requires integration with Microsoft Volume Shadow Copy Service, VSS.)
4. Create the necessary Snapshot copies on the storage systems.
5. Resume operations.

The NetApp and SnapDrive advantage – aside from the fact that all this happens more or less automatically without the operator having to know anything about the underlying storage – is that the entire Snapshot copy creation process can be done in a few seconds, so operations aren't disrupted.

Once you have a consistent Snapshot copy on the storage system, you can keep it on primary storage in case you need to restore a file or even roll back to that point in time. You can also back the Snapshot copy up to tape if desired, use NetApp SnapVault® to back up the Snapshot copy to secondary disk, or use NetApp SnapMirror technology to make a copy for disaster recovery. SnapDrive integrates with SnapMirror and SnapVault to do this automatically, when configured to do so.

NetApp continues to enhance SnapDrive functionality. To learn more about recent SnapDrive features, NetApp CSE Nick Triantos describes [thin provisioning and space reclamation in SnapDrive for Windows 5.0](#) and [self-service storage management with SnapDrive for UNIX](#) in two excellent posts on his blog.

Application Support Streamlines Backup, Recovery, & More NetApp Tool: SnapManager

The SnapManager set of offerings currently supports five popular enterprise applications (Microsoft Exchange, SQL Server, and SharePoint®, plus Oracle and SAP), which coordinate activities at the application level while leveraging the underlying capabilities of SnapDrive to coordinate with the host OS and storage system. (Another feature in this issue highlights the [unique capabilities of SnapManager for SharePoint](#).)

For example, consider the backup process with SnapManager for SQL Server (SMSQL), which carries out the necessary steps on the application side while SnapDrive for Windows performs essentially the same steps as described in the previous section:

Version	Unique Features
SnapManager for Exchange	Backup verification, Single Mailbox Restore
SnapManager for SQL Server	Database migration, database clone to any SQL Server
SnapManager for SAP	Certification with SAP® BR*Tools v6.4 and v7.0
SnapManager for Oracle	Close integration with RMAN, RAC, and ASM, Oracle® Database clone
SnapManager for SharePoint	Granular restore, SharePoint-centric management

1. A DBA requests a backup of a database: MySQLdb.
2. SnapManager for SQL freezes MySQLdb.
3. SnapManager for SQL generates a list of the drives used by MySQLdb.
4. SnapManager for SQL passes the list to SnapDrive.
5. SnapDrive determines which LUNs the drives are attached to and the corresponding FlexVol volumes.
6. SnapDrive stops activity on those LUNs and flushes data cached in the host operating system.
7. SnapDrive performs the Snapshot copy.
8. SnapDrive resumes host operations and signals success to SnapManager for SQL.
9. SnapManager for SQL unfreezes the database.

Again, despite the seemingly complicated process, the whole thing happens in a matter of seconds and the DBA doesn't need to know anything about how data is organized on the storage system or how it is sent over the network. SnapManager and SnapDrive work together to enable the application administrator to get the job done without having to constantly involve the storage administrator.

Typical capabilities that the SnapManager tools provide for applications include backup, restore, cloning, and disaster recovery, as well as scheduling and retention.

Each tool typically has features that are specific to that application. For instance, Single Mailbox Restore is an option with SnapManager for Exchange, and SnapManager for Oracle has the ability to integrate and coordinate with RMAN, ASM, and other critical features of Oracle. A command line interface (CLI) makes it possible to write scripts to integrate SnapManager actions such as Snapshot copies as part of the normal operation of an application.

Customized Application Management beyond Popular Platforms NetApp Tool: SnapDrive (Command Line Interface)

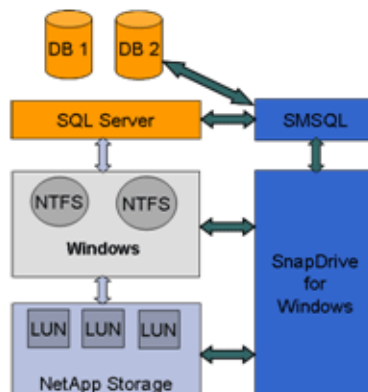


Figure 2) Database Backup Process.

Even though NetApp is committed to continue to add integration for popular software, we obviously will never be able to provide specific versions of SnapManager to cover all the myriad commercial applications in existence (or in-house applications that aren't built on SQL Server, Oracle, or some other database).

However, SnapDrive does include a command line interface `sdcli`, so that any application that can invoke a script can access SnapDrive functionality for backup, restore, etc. The following simple batch file is an example that one NetApp customer used to back up a database using SnapDrive for Windows. The database was put in hot backup mode and the script was called. When the script returns, hot backup mode is turned off.

```

REM add sdcli path - may need to edit
set path=%path%;C:\Program Files\NetApp\SnapDrive
REM ----- begin
REM change the drive letter as preferred after -d or -D
REM script maintains 7 snapshots, .1 = most recent .7 = oldest

IF EXIST \\FILERNAME\VOLSHARE\.snapshot\%computername%.daily.7 sdcli snap rename -d
e -o %computername%.daily.7 -n %computername%.daily.delete

IF EXIST \\FILERNAME\VOLSHARE\.snapshot\%computername%.daily.6 sdcli snap rename -d
e -o %computername%.daily.6 -n %computername%.daily.7
IF EXIST \\FILERNAME\VOLSHARE\.snapshot\%computername%.daily.5 sdcli snap rename -d
e -o %computername%.daily.5 -n %computername%.daily.6
IF EXIST \\FILERNAME\VOLSHARE\.snapshot\%computername%.daily.4 sdcli snap rename -d
e -o %computername%.daily.4 -n %computername%.daily.5
IF EXIST \\FILERNAME\VOLSHARE\.snapshot\%computername%.daily.3 sdcli snap rename -d
e -o %computername%.daily.3 -n %computername%.daily.4
IF EXIST \\FILERNAME\VOLSHARE\.snapshot\%computername%.daily.2 sdcli snap rename -d
e -o %computername%.daily.2 -n %computername%.daily.3
IF EXIST \\FILERNAME\VOLSHARE\.snapshot\%computername%.daily.1 sdcli snap rename -d
e -o %computername%.daily.1 -n %computername%.daily.2
REM set "-u = yes" for update of snapmirror, "-u = no" for no update

sdcli snap create -s %computername%.daily.1 -D e -u yes

IF EXIST \\FILERNAME\VOLSHARE\.snapshot\%computername%.daily.delete sdcli snap
delete -d e -s %computername%.daily.delete

REM ----- end
  
```

Get More from Your Application

If you need host-level storage integration or support for a homegrown app, SnapDrive by itself can provide an excellent solution that can simplify and accelerate many of the most common data management tasks. If you need to integrate your NetApp storage with Exchange, SQL Server, SAP, Oracle, or SharePoint, the SnapManager suite of tools can help you get the most value from your storage and your application. Using these tools, server and application administrators can perform important storage tasks without storage administrator assistance.