Technical Report

Playbook: Easily Assess Your Environment for NetApp Deduplication

Carlos Alvarez, NetApp
July 2011 | TR-3936

ABSTRACT

This technical report provides general guidance for assessing key storage attributes to quickly decide whether NetApp® deduplication is appropriate for a particular storage system. The guidance takes a conservative approach.

Additional testing and considerations can be applied for a more granular approach. See TR-3505, "NetApp Deduplication for FAS and V-Series Deployment and Implementation Guide," for details and best practices.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 OVERVIEW

Deduplication is a provisioning application option that you can enable on your storage nodes to eliminate duplicate data blocks and thereby reduce the amount of physical storage space used to store active data.

This document offers general guidance based on common practices that help you to quickly assess your environment to determine whether deduplication is appropriate. The guidance is conservative in nature to help you to achieve a successful implementation. A much more granular approach is also possible. See TR-3505, "Deduplication for FAS and V-Series Deployment and Implementation Guide," for details, considerations, and best practices.

The links to the technical reports referred to in this document are listed in section 8, "References."

1.1 REQUIREMENTS

Table 1) Deduplication requirements.

Requirement	Deduplication
Hardware	• NearStore® R200 • FAS2000 series • FAS3000 series • FAS3100 series • FAS3200 series • FAS6000 series • FAS6200 series • IBM N5000 series • IBM N7000 series **Note:** Starting with Data ONTAP® 7.3, the V-Series systems corresponding to the NetApp FAS systems and IBM N series Gateway systems listed above are also supported.
Minimum Data ONTAP version required	• Data ONTAP 7.3.x or later • (7-Mode for Data ONTAP 8.0.X only)
Licenses required	• Deduplication license • NearStore license (required for versions of Data ONTAP earlier than 8.0)
Volume type supported	• NetApp FlexVol® volumes only, no traditional volumes
Maximum volume size	• For Data ONTAP 8.0.1, the limit is 16TB for all platforms for deduplication. For earlier releases, see “Maximum Flexible Volume Size” in TR-3505.
Supported protocols	• All

1.2 CONSIDERATIONS FOR VMWARE AND VIRTUALIZATION

- Virtual machines (VMs) require special handling and configuration in order to work properly with a NAS or SAN system. This is especially true when running deduplication.

- VMs must be properly aligned in order to function properly. For information, see NetApp TR 3749, "NetApp and VMware vSphere Storage Best Practices."
- Lab Manager and VMware View™ Composer must not be deduplicated (as of March 1, 2011).

  These VMs use VMware® snapshots and these are all misaligned I/O. This situation causes a severe performance penalty and also has an exponentially negative effect on performance when deduplication is enabled.

2 HOW TO ASSESS YOUR SYSTEM FOR USING DEDUPLICATION

Before running deduplication in your environment, the storage controllers must be assessed for performance headroom to be able to accommodate the deduplication processing.

Based on some general numbers obtained from Operations Manager and/or Perfstat, the systems can be broken down into three general categories, as described in Table 2.

Table 2) System assessment.

Category	Description
Green	If the data type allows deduplication, green systems can generally have deduplication enabled to yield significant savings. • No greater than an average weekly CPU utilization of 50% ○ Cluster total CPU should be under 100% • No greater than any aggregate disk utilization of 50% • No greater latency on any LUN or volume than 10ms
Yellow	Yellow systems can have deduplication enabled given more careful consideration, such as data type, application response time demands, and user loads. If a controller has some “hot” volumes and some “cool” volumes, the cool volumes should have deduplication enabled first and then the overall health of the controller should be reassessed before considering the other hot volumes. • No greater than an average weekly CPU utilization of 75% ○ Cluster total CPU should be under 150% • No greater than any aggregate disk utilization of 75% • No greater latency on any LUN or volume than 20ms
Red	Red systems should generally not have deduplication enabled on them. However, if the CPU is low, and there are multiple aggregates, there is a possibility of enabling deduplication on a low utilized aggregate or a volume with low latency. The hot aggregates and volumes should not have deduplication enabled. In any circumstance, the controllers with CPU over a 75% average should not have deduplication enabled for any volume. • Any systems with an average weekly CPU utilization over 75% • Any systems with an aggregate disk utilization over 75% • Any systems with a LUN or volume latency over 20ms

3 METHODOLOGY FOR USING DEDUPLICATION

This section discusses how deduplication interacts with applications. Best Practice guides for the various types of applications are listed in section 8, "References." Consult these documents when building a solution.

3.1 GENERAL CONSIDERATIONS

The following guidelines are important when considering a storage controller for deduplication:

- Evaluate the performance characteristics of the system in a stable state by using data in Operations Manager and Performance Advisor. Perfstat is another option for gathering fine-grained data, such as the following:
  - Weekly % CPU utilization
  - Cluster total % CPU utilization
  - % aggregate disk utilization
  - LUN and volume latency (ms)
- Evaluate the data type on the system to see whether it is compatible with deduplication. See TR-3505, "NetApp Deduplication for FAS and V-Series Deployment and Implementation Guide."

For more reference documentation, see section 8, "References."

Best Practice
You can use the Space Savings Estimation Tool (SSET) to evaluate data sets, whether stored on NetApp or non NetApp storage, to estimate deduplication savings for unlisted data types.

- Evaluate the application's latency and the throughput requirements of any attached applications or servers.
- It is important to gather a baseline of performance data on the storage controller before enabling deduplication; this baseline can be referred to after deduplication has run to observe the performance impact.

3.2 VMWARE BEST PRACTICES

Reference documentation for using deduplication with VMware:

- TR 3747: Best Practices for File System Alignment in Virtual Environments
  http://media.netapp.com/documents/tr-3747.pdf
- TR 3428: NetApp and VMware Virtual Infrastructure 3 Storage Best Practices
  http://media.netapp.com/documents/tr-3428.pdf
- TR 3749: NetApp and VMware vSphere Storage Best Practices
  http://media.netapp.com/documents/tr-3749.pdf
- TR-3886: Understanding and Using vStorage APIs for Array Integration with NetApp Storage
  http://media.netapp.com/documents/tr-3886.pdf

VMware environments deduplicate extremely well. However, while working out the VMDK and datastore layouts, keep the following points in mind:

- Operating system VMDKs deduplicate extremely well because the binary files, patches, and drivers are highly redundant between virtual machines. To achieve maximum savings, keep these in the same volume.
- Application binary VMDKs deduplicate to varying degrees. Duplicate applications deduplicate very well; applications from the same vendor commonly have similar libraries installed and deduplicate somewhat successfully; and applications written by different vendors don't deduplicate at all.
- When deduplicated, application datasets have varying levels of space savings and performance impact based on application and intended use. Careful consideration is needed, just as with nonvirtualized environments, before deciding to keep the application data in a deduplicated volume.
- Transient and temporary data such as VM swap files, page files, and user and system temp directories do not deduplicate well and potentially add significant performance pressure when deduplicated. Therefore NetApp recommends keeping this data on a separate VMDK and volume that are not deduplicated. For more information on page files, see TR 3749, "NetApp and VMware vSphere Storage Best Practices"
- NetApp Data ONTAP 7.2.6 and 7.3.1 introduce a performance enhancement referred to as *intelligent cache.* Although it is applicable to many different environments, intelligent caching is particularly applicable to VM environments, where multiple blocks are set to zero as a result of system initialization. These zero blocks are all recognized as duplicates and are deduplicated very efficiently. The warm cache extension enhancement provides increased sequential read performance for such environments, where there are very large amounts of deduplicated blocks. Examples of sequential read applications that benefit from this performance enhancement include NDMP, NetApp SnapVault® software, and some NFS-based applications. This performance enhancement is also beneficial to the boot-up processes in VDI environments.
- The expectation is that about 30% space savings will be achieved overall. This is a conservative figure, and in some cases users have achieved savings of up to 80%. The major factor that affects this percentage is the amount of application data. New installations typically deduplicate extremely well, because they do not contain a significant amount of application data.

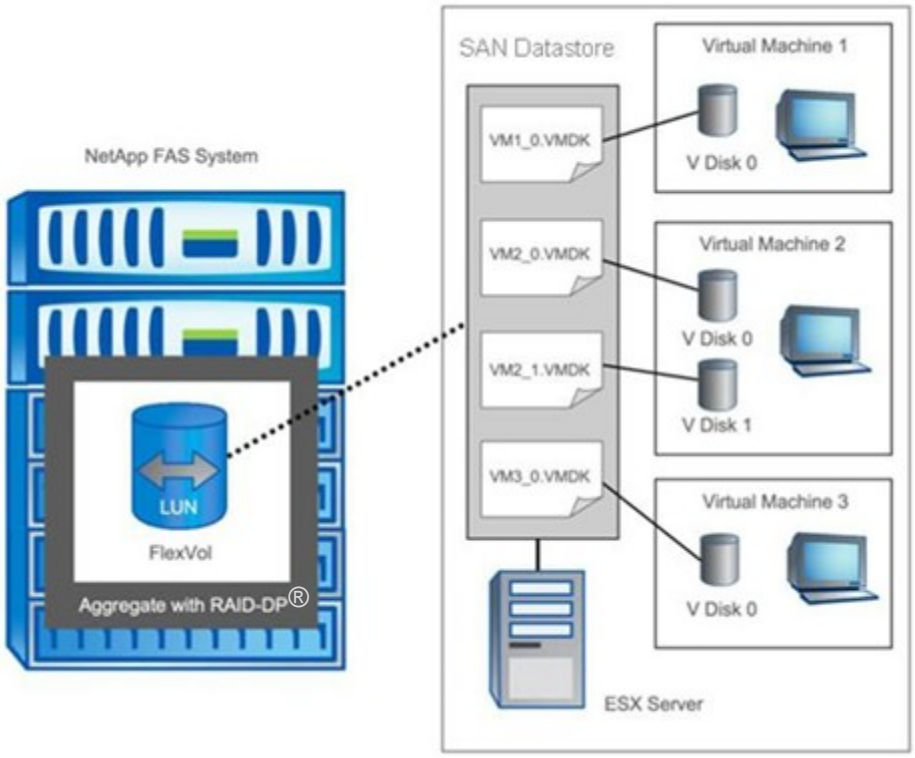
Important: In VMware environments, the need for proper partitioning and alignment of the VMDKs is extremely important (not just for deduplication). VMware must be configured so that the VMDKs are aligned on NetApp WAFL® (Write Anywhere File Layout) 4K block boundaries as part of a standard VMware implementation. To learn how to prevent the negative performance impact of LUN-VMDK misalignment, read TR-3747: "Best Practices for File System Alignment in Virtual Environments," TR-3428: "NetApp and VMware Best Practices Guide," or TR-3749: "NetApp and VMware vSphere Storage Best Practices." Also note that the applications in which performance is heavily affected by deduplication (when these applications are run without VMware) are likely to suffer the same performance impact from deduplication when they are run with VMware.

A deduplication and VMware solution on NFS is easy and straightforward. Combining deduplication and VMware with LUNs requires a bit more work.

VMFS DATASTORE ON FIBRE CHANNEL OR ISCSI—SINGLE LUN

Figure 1 shows the default configuration; this is the way a large number of VMware installations are done today. Deduplication occurs across the numerous VMDKs.

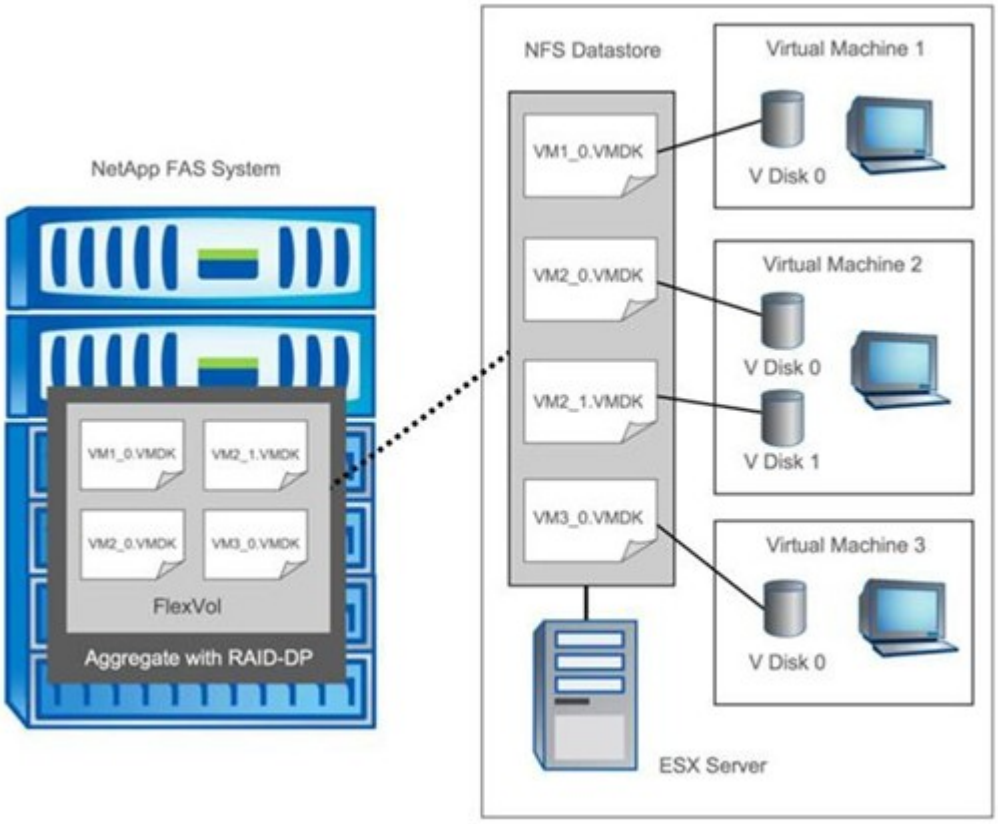
Figure 1) VMFS datastore on Fibre Channel iSCSI—single LUN.



VMWARE VIRTUAL DISKS OVER NFS OR CIFS

The newer configuration, available with VMware 3.0 and later, is easy to configure and allows deduplication to provide the most space savings.

Figure 2) VMware virtual disks over NFS or CIFS.



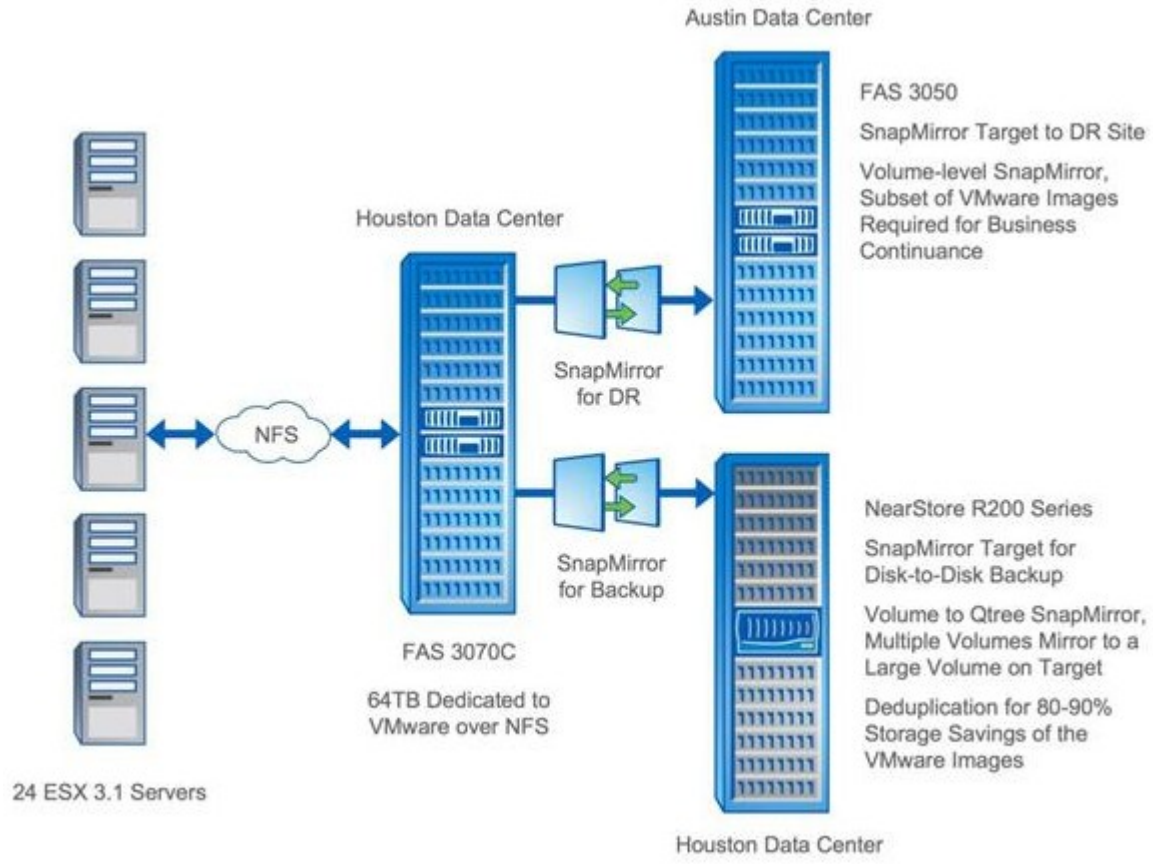
DEDUPLICATION ARCHIVE OF VMWARE

Figure 3 shows how deduplication has proven very useful in VMware archive environments.

Specifications for the example shown in Figure 3:

- VMware is configured to run over NFS.
- Uses approximately 1,800 clone copies of the master VMware image. These images are used to create virtual machines for primary applications and for test and development purposes.
- All 1,800 clone copies (~32TB) are stored on a FAS3070 in the Houston data center.
- The data is mirrored to the remote site in Austin for disaster recovery.
- Once per hour, the FAS3070 images are transferred to a NetApp NearStore R200 by using NetApp SnapMirror®.
- Deduplication is run nightly on the NearStore R200, and the VMware images are reduced in size by 80% to 90%.

Figure 3) Archive of VMware with deduplication.



3.3 MICROSOFT EXCHANGE SERVER BEST PRACTICES

Reference documentation for using deduplication with Microsoft Exchange:

- TR 3578: Microsoft Exchange Server 2007 Best Practices Guide
  http://media.netapp.com/documents/tr-3578.pdf
- TR 3824: Storage Efficiency and Best Practices for Microsoft Exchange Server 2010
  http://media.netapp.com/documents/tr-3824.pdf

If Microsoft® Exchange and NetApp deduplication are used together, consider the following points:

- In some Exchange environments, extents are enabled to improve the performance of database validation. Enabling extents does not rearrange blocks on disk that are shared between files by deduplication on deduplicated volumes. Enabling extents does not predictably optimize sequential data block layout when used on deduplicated volumes, so there is no reason to enable extents on deduplicated volumes.
- Beginning with Microsoft Exchange 2010, single instancing storage is no longer available. NetApp deduplication for FAS and V-Series offers significant savings for primary storage running Exchange 2010. Similar savings can be expected with Microsoft Exchange 2007 if Exchange single instancing is disabled. NetApp recommends running SSET on your environment to estimate the deduplication savings that your environment can achieve.

3.4 MICROSOFT SQL SERVER BEST PRACTICES

Deduplication can provide space savings in Microsoft SQL Server® environments, but proper testing should be done to determine the savings for your environment. Use the Space Savings Estimation Tool (SSET 3.0) to estimate the amount of savings that would be achieved with deduplication.

Our testing has shown deduplication savings in the range of 15%. NetApp recommends running SSET against your data to estimate the deduplication savings that your environment can achieve.

A Microsoft SQL Server database uses 8K block size. Each SQL Server block is composed of two NetApp 4K blocks. Recall that deduplication works at the 4K block level. So, although Microsoft SQL Server places a unique header at the beginning of each SQL Server block (corresponding to one 4K block), the rest of the blocks might still contain duplicates. This means that deduplication might provide significant savings when comparing the 4K blocks within the volume.

3.5 MICROSOFT SHAREPOINT BEST PRACTICES

If Microsoft SharePoint® and NetApp deduplication for FAS are used together, consider the following points.

Make sure that space is available in the volume before using the `sis on` command. If this command is used on a flexible volume that already has data and is completely full, it fails. Up to 6% of the total data size is needed for deduplication of metadata files.

Deduplication is transparent to SharePoint. The block-level changes are not recognized by SharePoint, so the SharePoint database remains unchanged in size, even though there are capacity savings at the volume level.

3.6 ORACLE BEST PRACTICES

Deduplication can provide savings in Oracle® environments, but proper testing should be done to determine the savings for your environment. You can use SSET 3.0 to estimate the amount of savings that would be achieved with deduplication.

Deduplication savings depend on the Oracle configurations.

A typical Oracle data warehouse or data mining database uses 16K or 32K block size, made of multiple NetApp 4K blocks. Although Oracle places a unique identifier at the beginning of each Oracle block, and a near-unique identifier at the end of each Oracle block, the rest of the blocks within each Oracle block might still contain duplicates. This means that deduplication might provide significant savings when comparing the 4K blocks within the volume.

Oracle OLTP databases typically use an 8K block size. Oracle once again places a unique identifier at the beginning and end of each Oracle block, making the 4K blocks unique. Testing has shown that these environments typically do not have a significant number of duplicate blocks and do not show very much deduplication savings.

One additional case to consider is when a table space is created or extended. In this case Oracle initializes the blocks and commits many of them in the same transaction. This results in the creation of duplicate blocks, allowing deduplication to provide savings. However, as these blocks are filled with incoming data the space savings shrink.

3.7 SYMANTEC BACKUP EXEC BEST PRACTICES

If Symantec® Backup Exec™ and NetApp deduplication for FAS are used together, consider the following point: Deduplication savings with Backup Exec are not optimal because Backup Exec does not block-align data when it writes files out to its volumes. The net result is that there are fewer duplicate blocks available to deduplicate.

3.8 BACKUP BEST PRACTICES

There are various ways to back up your data. Here are a few deduplication considerations for your backup environment:

- For optimal backup throughput, it is a best practice to make sure that:
  - Deduplication operations initiate only after your backup completes
  - Deduplication operations on the destination volume complete before initiating the next backup
- When backing up data from multiple volumes to a single volume, you might achieve significant space savings from deduplication beyond just the deduplication savings from the source volumes. This is because you are able to run deduplication on the destination volume, which could contain duplicate data from multiple source volumes.
- If you are backing up data from your backup disk to tape, consider using SMTape to preserve the deduplication savings. Using NDMP to tape does not preserve the deduplication savings on tape.

4 SCHEDULING

Scheduling is a critical aspect of running deduplication successfully. On most storage controllers, up to eight simultaneous deduplication processes can be run at once. The best time to schedule the deduplication processes to run is the time of lowest user I/O activity on the controller. The default times are set to run on Saturday and Sunday morning every week. The schedule can be fine-tuned to run at other off-peak times. NetApp recommends staggering the processes so they have time to complete and are not competing for CPU and disk resources. When running multiple deduplication processes simultaneously, if possible, it is preferable to have different aggregates running deduplication simultaneously rather than within a single aggregate.

To save network bandwidth, it may be desirable to schedule SnapMirror processes after the deduplication processes on the primary storage systems complete.

Best Practice
When using SnapMirror and SnapVault, deduplication should complete on the primary system. SnapVault automatically starts deduplication on the destination (if enabled). Scheduling cannot be controlled manually on the destination system.

If at any time there are eight processes running, any new process will be queued until there is an open slot. (For a description of queuing and restarting, see TR-3505, "Deduplication for FAS and V-Series Deployment and Implementation Guide.")

When the deduplication processing is happening, user I/O takes priority over the background deduplication process.

5 SETTING ALERTS AND ALARMS

When monitoring deduplication performance after enabling it on a storage controller, you should consider a few factors for setting up alerts and alarms.

During deduplication processing, the storage controller may experience 100% CPU utilization periods and high (>90%) disk utilization, especially if multiple deduplication processes are scheduled simultaneously. Recall that deduplication is a “nice” process, and gives priority to other applications running on the storage systems; however, if resources are unused, deduplication attempts to use them.

Therefore, when configuring alarms on a deduplication enabled system, it is important to set up “combo” alarms. High CPU or high disk utilization alone may not be a problem and may not affect user or application performance.

Protocol, volume, and LUN latency is the best way to determine what response time the end user or application is getting. Alarms should be configured to take this into account.

Example 5-1) Alarms.

- CPU > 90% AND Disk > 90% AND Latency > 20ms
- Disk > 90% OR CPU > 90% AND Latency > 20ms

These thresholds vary depending on the application, OS, and end-user requirements. A database requires latency under 20ms, a VM might be fine up to 50ms, and a NAS share might be fine up to 100ms.

6 MONITORING

To achieve continued reliable and efficient operation, the storage controllers must be monitored with Operations Manager and periodically assessed to see if deduplication is working properly and is not causing performance problems.

There are three main system parameters to monitor: CPU, disk I/O, and latency. Table 3 provides guidance for evaluating these parameters.

Table 3) Evaluating system parameters.

Category	Description
CPU	• CPU on a healthy storage controller should remain below 75%. • CPU on a healthy cluster should remain below 150% total. • Peak CPU does not necessarily indicate a problem. • Average CPU over the entire week should be used as a baseline. • CPU can potentially spike on the weekend when media scrubs happen. • CPU may spike up to 100% when deduplication processing is run. • CPU may spike up hourly or nightly when Snapshot®, SnapMirror, and SnapVault transfers take place. • CPU may not correlate to user I/O (net/FCP in/out).
Disk I/O	• Disk utilization in a healthy aggregate should remain under 75%. • Disk utilization may spike during media scrubs, deduplication processes, and Snap operations. • Average disk utilization over the entire week should be used as a baseline. • Weekly average disk utilization should correlate to user I/O for volumes and LUNs in that aggregate. • If there is a RAID rebuild event such as a failed disk, disk utilization spikes up to 100%. If disk utilization is at 100% with no explanation, check system RAID status (`aggr status -r`). • Aggregates with disk utilization over 75% should not have deduplication enabled on any volume in them.
Latency	• Latency is measured per protocol at the system and volume levels. • Overall controller latency per protocol is a good baseline. • Individual volume and LUN latency is a good measure to see whether deduplication can be enabled on that specific volume. • Volumes or LUNs with latency over 20ms should not have deduplication enabled even if the overall controller latency is low.

7 APPENDIXES

7.1 DEDUPLICATION COMMANDS

This appendix lists the various deduplication commands.

Table 4) Deduplication commands.

Command	Summary
sis on <vol/volname >	Enables deduplication on the flexible volume specified.
sis start <vol/volname >	Begins the deduplication process on the flexible volume specified.
sis start –s <vol/volname >	Begins the deduplication process on the flexible volume specified and performs a scan of the flexible volume to process existing data. This option is typically used at initial configuration and deduplication on an existing flexible volume that contains undeduplicated data. (There is no need to use this option on a volume that has just been created and doesn't contain any data.)
sis start –sp <vol/volname >	Begins the deduplication process on the flexible volume specified by using the existing checkpoint information, regardless of the age of the checkpoint information. This option should only be used with the –s option.
sis start –d <vol/volname >	Deletes the existing checkpoint information. This option is used to delete checkpoint information that is still considered valid. By default, checkpoint information is considered invalid after 24 hours.
sis status [-l] <vol/volname >	Returns the current status of deduplication for the specified flexible volume. The –l option displays a long list.
sis config [-s sched] <vol/volname >	Creates an automated deduplication schedule. When deduplication is first enabled on a flexible volume, a default schedule is configured, running it each day of the week at midnight. If the auto option is used, deduplication is triggered when 20% new data is written to the volume. Starting with Data ONTAP 7.3.1, the 20% threshold can be adjusted by using the auto@num option, where num is a two-digit number to specify the percentage.
sis stop <vol/volname >	Suspends an active deduplication process on the flexible volume.
sis off <vol/volname >	Deactivates deduplication on the flexible volume specified. This means that there will be no additional change logging or deduplication operations, but the flexible volume remains a deduplicated volume, and the storage savings are preserved. If this command is used, and then deduplication is turned back on for this flexible volume, the flexible volume must be rescanned with the sis start –s command.
sis check <vol/volname >	Verifies and updates the fingerprint database for the flexible volume specified; includes purging stale fingerprints (diag mode only).
sis stat <vol/volname >	Displays the statistics of flexible volumes that have deduplication enabled (diag mode only).
sis undo <vol/volname >	Reverts a deduplicated volume to a normal flexible volume (diag or advanced

	mode only).
sis revert_to <ver> <vol/volname>	Converts the sis metafiles to appropriate lower version format (currently 7.2 or 7.3). **Note:** When no volume name is provided, revert_to runs on all volumes that have deduplication enabled.

7.2 TYPICAL DEDUPLICATION SPACE SAVINGS BY DATA TYPE

This appendix provides typical deduplication space savings by data type and describes some key factors that affect the savings.

Comprehensive testing with various datasets has been performed to determine typical space savings in different environments. These results were obtained in three ways:

- Running deduplication on various production datasets in NetApp
- NetApp systems deployed in the real world running deduplication
- NetApp personnel and end users running a simulation tool on various datasets

Table 5 summarizes typical deduplication space savings by data type.

Table 5) Typical deduplication space savings by data type.

Data Type	Application Type	Deduplication Savings
File services/IT infrastructure		30%
Virtual servers and desktops		70%
Database	Oracle OLTP	0%
	Oracle DW	15%
	SQL Server	20%
E-mail, collaborative	Exchange 2003/2007	3%
	Exchange 2010	15%
Engineering data		30%
Geoseismic data		3%
Archival data		25%
Backup data		95%

These results are based on internal testing and customer feedback and are considered realistic and typically achievable. Savings estimates can be validated for an existing environment by using the Space Savings Estimation Tool (SSET).

Note the following information:

- Nonrepeating archival data such as image files and encrypted data is generally not considered a good candidate for deduplication.
- The deduplication space savings in this table result from deduplicating a dataset one time, with the following exception. In cases where the data is being backed up or archived over and over again, the realized storage savings get better and better, achieving 20:1 (95%) in many instances. The backup case also assumes that the backup application is maintaining data coherency with the original, and that the data's block alignment will not be changed during the backup process. If these criteria are not true, then there can be a discrepancy between the space savings recognized on the primary and secondary systems.

8 REFERENCES

The documents listed in this section enhance and broaden your knowledge of the information presented in this document. They are available on the NetApp Support site (formerly NOW™): http://now.netapp.com/

- TR 3505: NetApp Deduplication for FAS and V-Series Deployment and Implementation Guide
  http://media.netapp.com/documents/tr-3505.pdf
- TR 3578: Microsoft Exchange Server 2007 Best Practices Guide
  http://media.netapp.com/documents/tr-3578.pdf
- TR 3824: Storage Efficiency and Best Practices for Microsoft Exchange Server 2010
  http://media.netapp.com/documents/tr-3824.pdf
- TR-3749: NetApp and VMware vSphere Storage Best Practices
  http://media.netapp.com/documents/tr-3749.pdf
- TR-3712: Oracle VM and NetApp Storage Best Practices Guide
  http://media.netapp.com/documents/tr-3712.pdf
- TR 3710: Operations Manager, Provisioning Manager, and Protection Manager Best Practices Guide
  http://media.netapp.com/documents/tr-3710.pdf
- TR 3747: Best Practices for File System Alignment in Virtual Environments
  http://www.netapp.com/us/library/technical-reports/tr-3747.html
- TR 3748: NetApp and VMware Virtual Infrastructure 3 Storage Best Practices
  http://media.netapp.com/documents/tr-3428.pdf
- TR 3749: NetApp and VMware vSphere Storage Best Practices
  http://www.netapp.com/us/library/technical-reports/tr-3749.html
- TR-3886: Understanding and Using vStorage APIs for Array Integration with NetApp Storage
  http://media.netapp.com/documents/tr-3886.pdf

9 VERSION TRACKING

Version 1	June 2011	Initial release

NetApp
www.netapp.com