White Paper

EMC VNXe3200 File Deduplication and Compression
A Detailed Review

Abstract

This white paper describes EMC® VNXe™ File Deduplication and Compression for the VNXe3200 storage system. This feature increases the efficiency with which network-attached storage (NAS) data is stored. This paper outlines the motivation for increased storage efficiency, how File Deduplication and Compression works, the principles, and technical architecture behind it.

February 2015

EMC²

Table of Contents

Executive summary

As storage usage increases exponentially, improving storage efficiency is critical for many customers today. There are many viable solutions to achieve this objective. The solution you implement is based on what you believe is a suitable or effective approach to achieve storage efficiency in your production environment. If you are looking to reduce the cost of storing data at the file-level without affecting the end-user experience, the EMC® VNXe™ File Deduplication and Compression feature should be considered.

The VNXe File Deduplication and Compression feature increases file storage efficiency by eliminating redundant data from the files stored in file-level storage, thereby saving storage space and cost. File-level deduplication gives the Storage Processor (SP) the ability to process files in order to compress them, and the ability to share the same instance of the data when the files are identical. This feature is applied to file systems and VMware NFS datastores on the VNXe3200 storage system, and is transparent to access protocols.

This whitepaper introduces the VNXe File Deduplication and Compression's product design, which offers ease of management and intelligent storage efficiency. This paper illustrates the concepts of File Deduplication and Compression and shows how you can manage deduplication on the VNXe3200 storage system. The feature is designed to be a simple but powerful addition to an already strong portfolio of VNXe features.

Audience

This white paper is intended for use by EMC field personnel and customers who are familiar with EMC VNXe technology but are not familiar with VNXe File Deduplication and Compression.

Terminology

Common Internet File System (CIFS) – An access protocol that allows users to access files and folders from Windows hosts located on a network. User authentication is maintained through Active Directory and file access is determined by directory access controls.

Deduplication – The process used to compress redundant data, allowing space to be saved on a storage resource. When multiple files have identical data, the storage resource stores only one copy of the data, and shares that data between the multiple files.

File System – A storage resource that is associated with file access protocols. File systems are accessed through Windows shares (CIFS protocol) or Linux/UNIX shares (NFS protocol).

Network-Attached Storage (NAS) Server – A file-level storage server that provides access to multiple clients. A NAS Server is required in order to create file systems that contain CIFS or NFS shares, or VMware NFS datastores.

Network File System (NFS) – An access protocol that allows users to access files and folders from Linux/UNIX hosts located on a network.

Network Data Management Protocol (NDMP) – Protocol that provides a standard for backing up file servers on a network. NDMP allows centralized applications to back up file servers running on various platforms and platform versions. NDMP reduces network congestion by isolating control path traffic from data path traffic, which permits centrally managed and monitored local backup operations.

Portable Archive Interchange (PAX) – An AIEEE/Posix standard archive protocol that works with standard UNIX tape formats and provides file-level backup and recovery operations.

Storage resource – An addressable and configurable storage instance associated with a specific quantity of storage, storage server, storage pool, and protocol. Types of storage resources include LUNs, VMware datastores, and file systems.

Storage Processor (SP) – A hardware component that provides the processing resources for performing storage operations such as creating, managing, and monitoring storage resources.

Reduplication – Process to undo the effect of deduplication on a file. If the file was compressed, it will be decompressed. If there are multiple instances of the file data, then a copy of the file data is made so that blocks are not shared between instances of the file.

Unisphere – A web-based management environment for creating storage resources, configuring and scheduling protection for stored data, and managing and monitoring other storage operations on the VNXe3200 system.

Unisphere Command Line Interface (CLI) – An interface that allows a user to perform storage-related tasks by using commands typed into a command line.

Detailed overview

Data reduction technology choices

There are a number of technologies classified under data reduction or deduplication. Two major data reduction technologies utilized in the VNXe3200 storage system are:

- **File-level deduplication** - Provides relatively modest space savings. It does not require substantial CPU and memory resources to implement.

- **Compression** - Often regarded as a separate mechanism from deduplication, compression can also be viewed as variable, bit-level, intra-object deduplication. Compression alters the way data is stored, mainly to improve storage efficiency. In fact, it offers by far the greatest space savings of all the techniques listed for typical NAS data, and is relatively modest in terms of its resource footprint.

VNXe File Deduplication and Compression combines file-level deduplication and data compression technologies to provide maximum benefit for the resources that are required to provide storage efficiency, while still minimizing the client impact on mission-critical files. This feature processes file data, not metadata. If multiple files contain the same data but different names, the files are deduplicated. Deduplication eligible files can also have different permissions and timestamps.

For example, if there are 50 unique files that can be deduplicated in a file system, 50 unique files will still exist but the data will be compressed, yielding space savings. If there are 70 identical copies of a presentation document in the file system, 70 files will still exist but they will all share the same file data. In this example, the data usage will decrease by a factor of almost 70. In addition, the one instance of the file data shared by the 70 files will also be compressed, providing further space savings.

Minimizing client impact

The VNXe3200 storage system performs all deduplication processing as a background asynchronous operation that acts on file data after it is written into the file-level storage resource. It does not process data while file data is being written into the storage resource. This avoids latency on the client side as access to production data is sensitive to latency.

In addition to processing data in the background, File Deduplication and Compression avoids processing data that is currently considered "active" or "hot". This is defined by how recently clients have accessed or modified the files. Typically only a small proportion of the data in production storage is in active use. This means that the bulk of the data is processed in storage without affecting the production workload.

The policy engine uses a pre-defined default policy to scan the files in a deduplication-enabled storage resource. This is based on an analysis of how typical files age from active use to disuse in various company settings of different industry and sector types.

Targeting Files

VNXe File Deduplication and Compression targets older files and avoids new files that are considered active. Therefore, this feature scans each file-level storage resource that is marked to be processed no more than once every 7 days. Administrators can prompt the system to scan a specific storage resource immediately, if required.

File size is also a criterion to decide whether a file is a candidate for deduplication. VNXe File Deduplication and Compression scans a file for its minimum and maximum size to determine if it should be deduplicated.

Certain file types compress well, such as text files (.doc), whereas other file types already contain compressed data such as media files (.mpg) and image files (.jpg). File systems used for home directories and office file distribution tend to achieve higher space savings whereas file systems that contain a high percentage of image files and software patch updates usually achieve lower percentage of space savings with respect to the logical file system data size.

VNXe File Deduplication and Compression scans files based on system-defined policy criteria (see Table 1 for more details). However, there are two policy settings that can be configurable by the end user. Administrators can define filters for the VNXe3200 storage system to avoid processing files based on the file extension. They can also define filters based on pathname to restrict directories from being scanned.

Table 1 VNXe File Deduplication and Compression Policy Criteria

Setting	Definition	Default Value
Access Time	Length of time in days the file has not been accessed.	15 days
Modification Time	Length of time in days the file has not been modified.	15 days
Minimum Size	Files less than this size will not be deduplicated.	24 KB
Maximum Size	Files greater than this size will not be deduplicated.	8 TB
Detection Method	Hash used to detect duplicate files.	SHA1
Case Sensitive	Defines if case-sensitive (NFS environments) or case-insensitive (CIFS environments) string comparisons will be used.	Off
File Extension Exclude List*	Files with the specified extensions will not be deduplicated.	None
Pathname Exclude List*	Directories with the specified pathname will not be deduplicated.	No pathnames
Minimum Scan Interval	Frequency with which the deduplication policy engine will scan a deduplication-enabled file system.	7 days
CPU % High Water Mark	If the CPU reaches this level, deduplication will throttle down.	90%
CPU % Low Water Mark	If deduplication is throttled down and the CPU level returns to this level, deduplication will throttle back up.	40%
CIFS Compression Attribute	Enables CIFS deduplication.	Enabled
Backup Data High Water Mark	For NDMP to back up, the space-reduced size must reach this level of the logical file size.	90%

* Settings are configurable in Unisphere

CPU Throttling

The impact of deduplication processing on the VNXe3200 storage system is controlled through an automated scheduling mechanism and is self-regulating based on the CPU load. Each SP scans and deduplicates only one file-level storage resource at a time. While scanning and deduplicating files, the CPU load of the SP is monitored. If the VNXe3200 storage system detects that the CPU load exceeds a system-defined threshold, activity is throttled down until the CPU load has decreased to less than a low-activity threshold. This means that the deduplication and reduplication processes effectively consume CPU cycles that will otherwise be idle and avoid affecting the system's ability to satisfy client activity.

CIFS Compression

The CIFS compression attribute feature allows Windows CIFS clients to manually deduplicate and reduplicate individual files or directories that do not meet the defined policy. Windows Explorer shows the names of deduplicated files marked in a different color (the default is blue). The VNXe3200 storage system attempts to preserve a deduplicated file in the deduplicated state even when it is written to. As soon as deduplication is disabled on the VNXe3200 storage system, CIFS compression is no longer allowed until the administrator enables or suspends deduplication.

The VNXe3200 storage system shows deduplicated files as being compressed to the CIFS client. As shown in Figure 1, CIFS clients can deduplicate and reduplicate individual files and directories using the Windows Explorer advanced attribute, "Compress contents to save disk space" (Right-click a non-deduplicated file, and then click **Properties** › **Advanced** › **Compress contents to save disk space**). Marking a directory affects only new files that are added to the directory. It does not affect existing files in the directory.
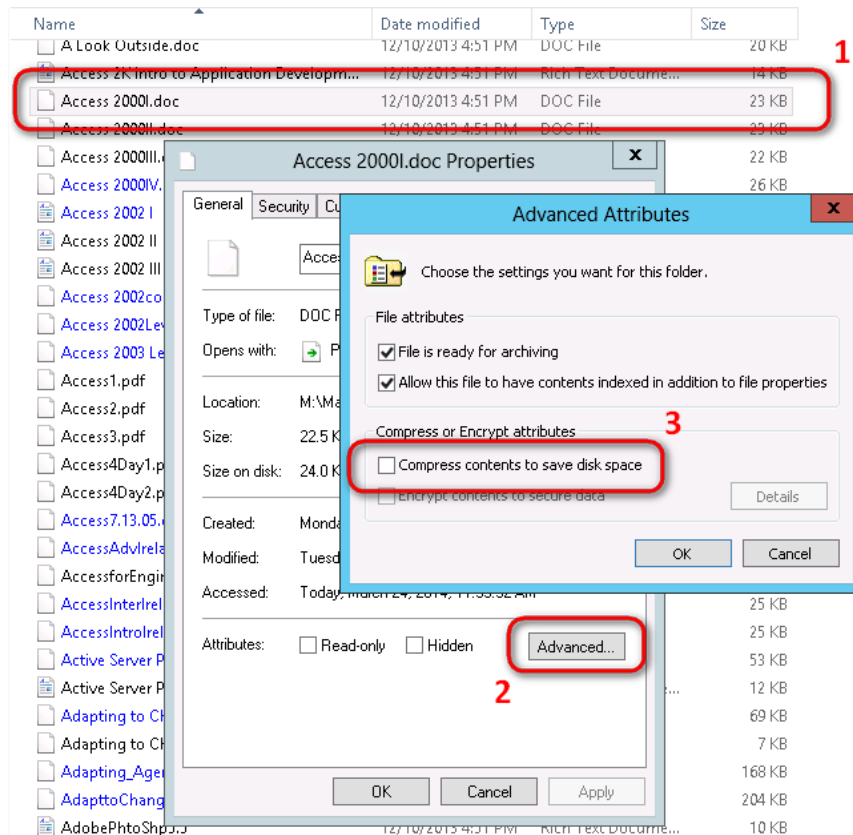

Figure 1: Manually compressing a file

The space reduction process

VNXe File Deduplication and Compression has a policy engine that specifies data for exclusion from processing on file systems or VMware NFS Datastores, and decides whether to deduplicate specific files based on the policy criteria. Figure 2 depicts a typical file system containing files with different attributes (i.e. last modified date, file size, and type).
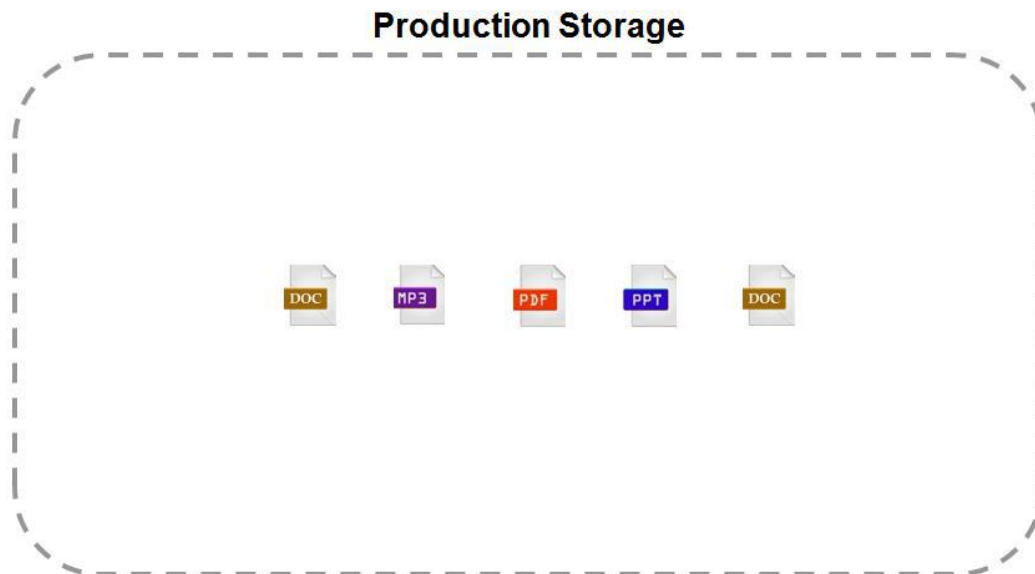
Production Storage

Figure 2: File system contains various files

When VNXe File Deduplication and Compression is enabled, it dynamically allocates a "hidden store", which holds deduplicated data and the hash table used for single instancing. The policy engine then periodically scans for files that match the policy criteria, as shown in Figure 3.

Production Storage   **Hidden Store**

Hash Table

Policy Engine

1. Eligible? (policy check)
2. Copy, compress, and hash.
3. Redundant? (hash check)
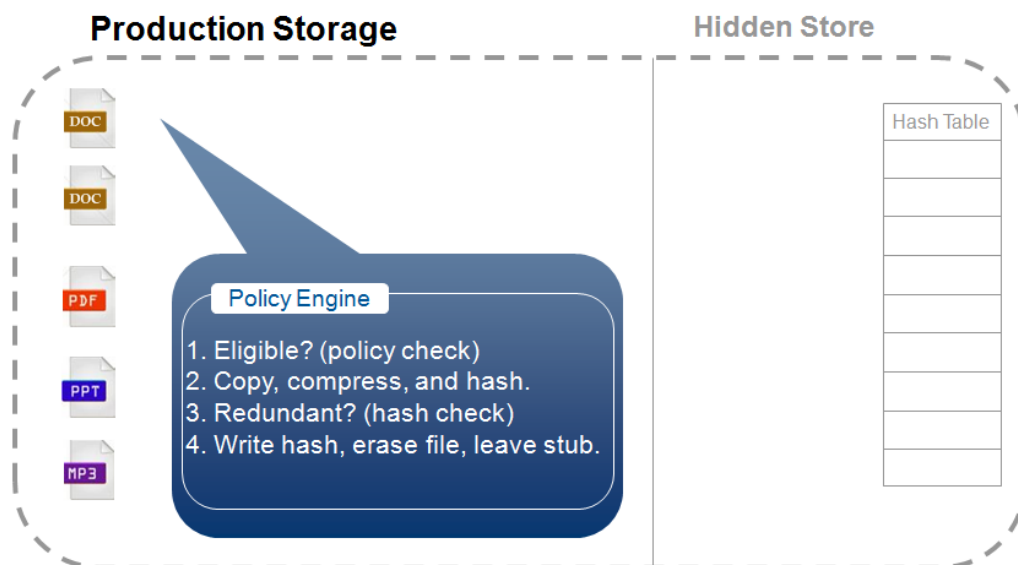4. Write hash, erase file, leave stub.

Figure 3: Scanning a file against the policy criteria

If the file meets the policy criteria, it is copied to the hidden store for compression, and then it is hashed. The hash is used to compare with other hashes in the hash table to determine if there is another copy the policy engine has processed before. In this example it is a unique hash. So, the hash is stored in the hash table for later comparisons, as shown in Figure 4.

Production Storage　　**Hidden Store**

DOC STUB

PP1FED81

Hash Table
PP1FED81

Policy Engine
1. Eligible? (policy check)　　**YES**
2. Copy, compress, and hash.
3. Redundant? (hash check)　　**NO**
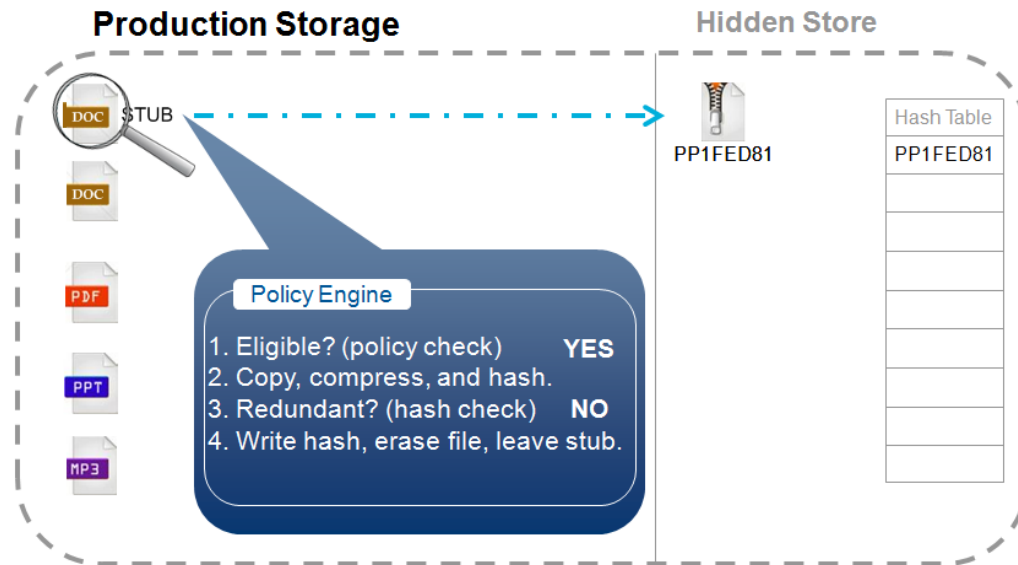4. Write hash, erase file, leave stub.

Figure 4: File meets policy criteria and gets processed

The space that the file data occupied in the user portion of the file system is freed and the file's internal metadata is updated to reference the copy of the data in the hidden store as shown in Figure 5.

Production Storage　　**Hidden Store**

DOC STUB
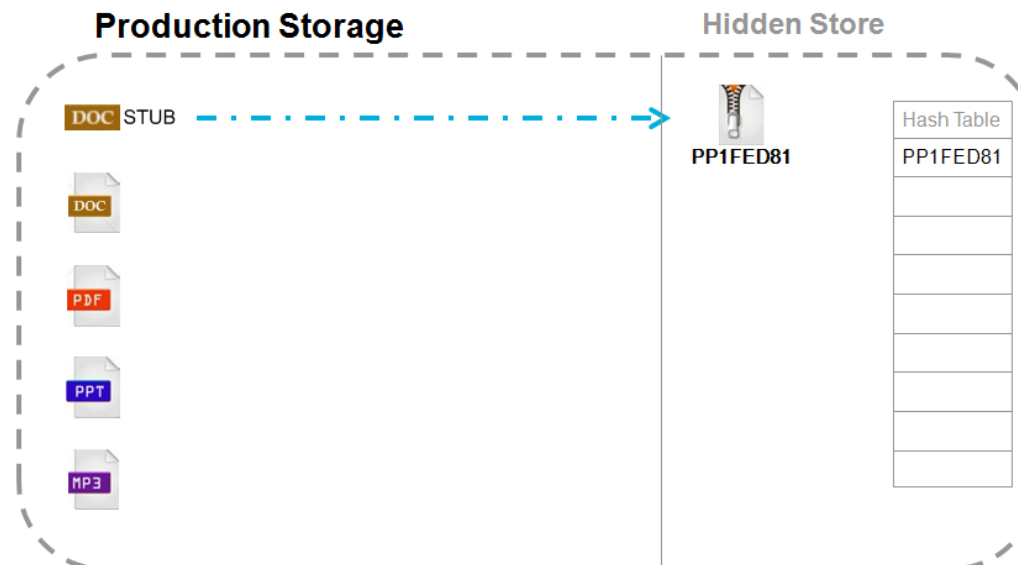
PP1FED81

Hash Table
PP1FED81

Figure 5: Stub file points to the file data in the hidden store

If the data associated with the file has been identified before, the space it occupies is freed and the internal file metadata is updated to reference the instance already in the hidden store, as shown in Figure 6.

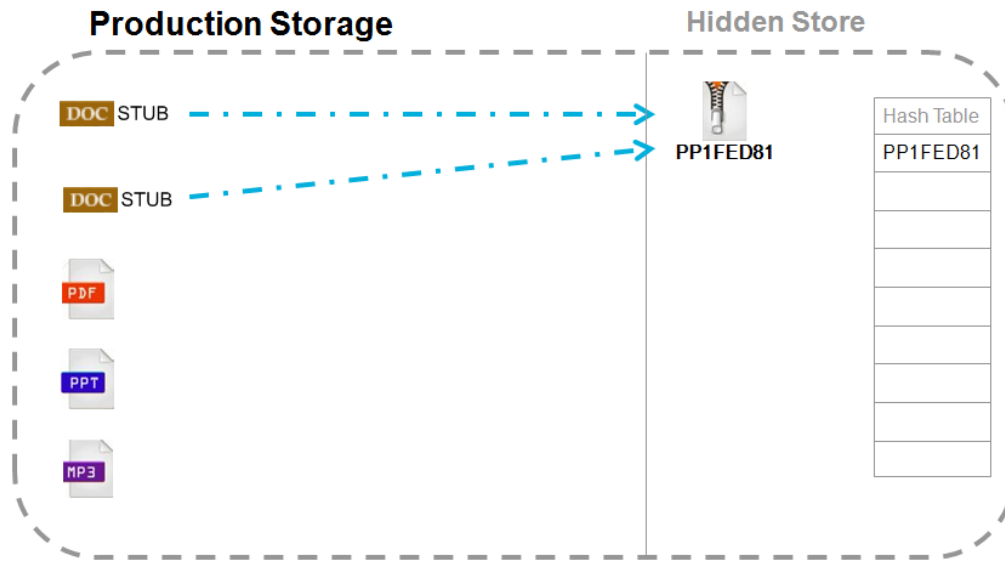Figure 6: A duplicate file is detected

Note: The VNXe3200 storage system detects non-compressible files, such as .mpg or .jpg files, and stores them in their original form. However, these files can still benefit from file-level deduplication.

Management

You can manage the VNXe File Deduplication and Compression feature through Unisphere or Unisphere CLI. For more information regarding Unisphere CLI, please refer to the Unisphere CLI User Guide available on EMC Online Support.

Enabling/Disabling Deduplication

VNXe File Deduplication and Compression can be enabled when provisioning file systems or VMware NFS Datastores. To enable this feature for a new file system, select the **Enable Deduplication and Compression** checkbox in the **File Systems Settings** wizard pane when creating a file system, as shown in Figure 7.
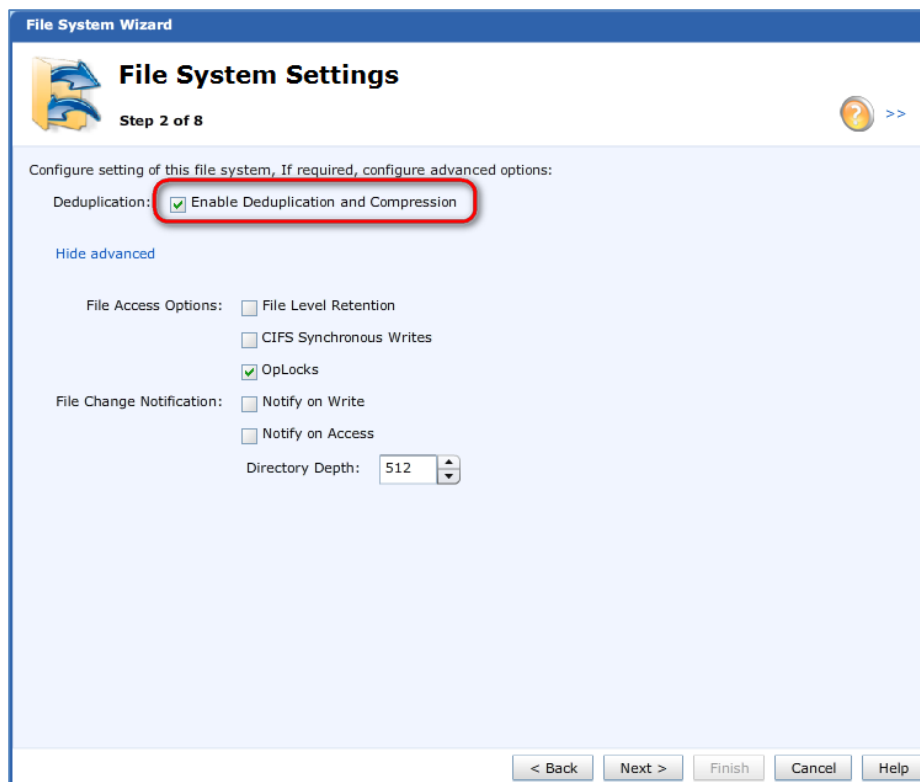

Figure 7: Enable during File System creation wizard

To enable this feature for a new VMware NFS datastore, select the **Enable Deduplication and Compression** checkbox in the **Specify Datastore Type** wizard pane when creating a VMware NFS datastore, as shown in Figure 8.
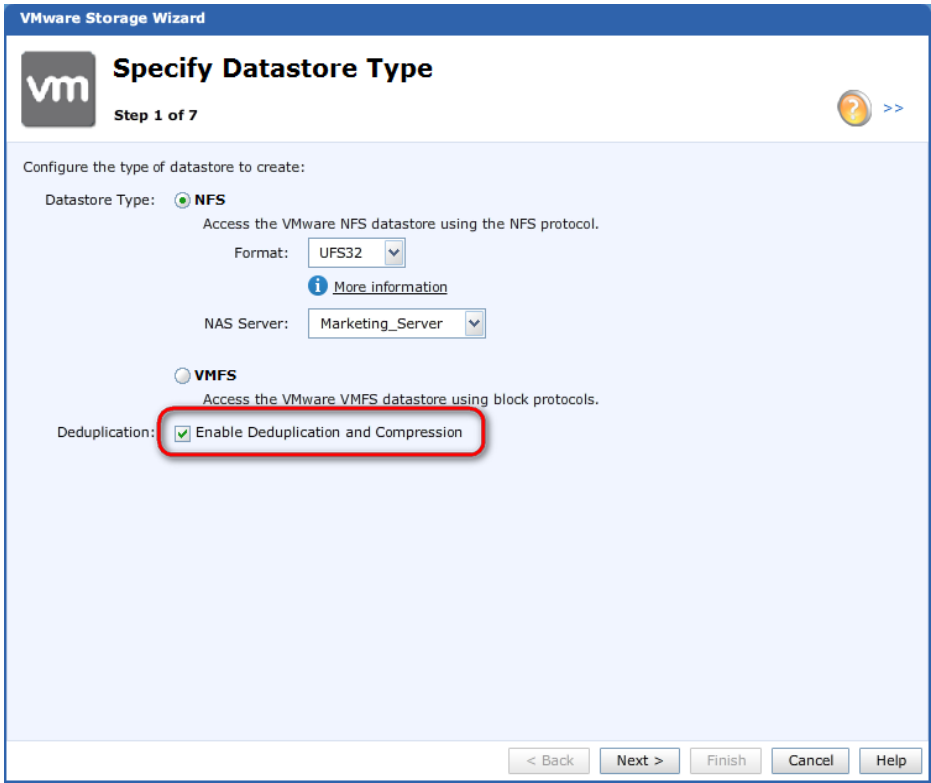
Figure 8: Enable during VMware NFS Datastore creation wizard

Alternatively, this feature can be enabled after the storage resource has been created by choosing the storage resource and selecting **Details** › **Deduplication** tab › **Enable deduplication,** as shown in Figure 9.

Figure 9: Enable feature for existing storage resource

Enabling this feature prompts an immediate scan of the storage resource. However, if another scan is already in progress on that particular NAS Server, it will not be initiated.

Once the VNXe File Deduplication and Compression feature is enabled, Windows CIFS clients can make use of the Windows Explorer advanced attribute, "Compress content to save disk space".
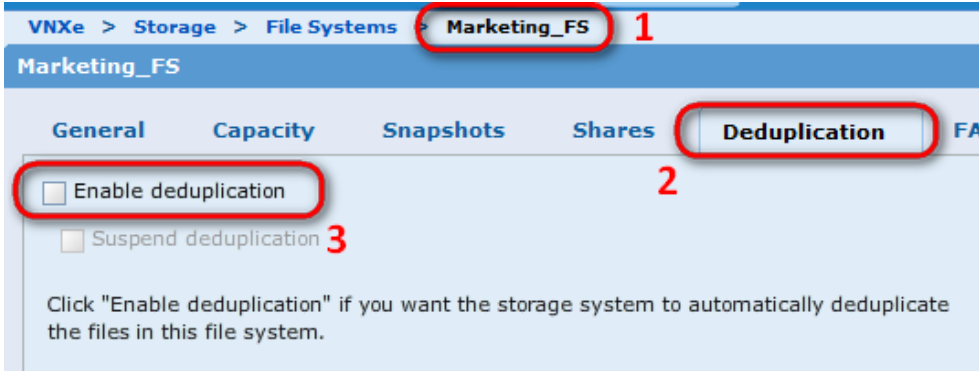
After an administrator enables a storage resource for deduplication, that storage resource is periodically scanned for more files to deduplicate. Administrators can view the state of the deduplication process in the Deduplication tab, located in the Details page for the storage resource. In the case that the Suspend deduplication setting is selected, the policy engine does not scan for any additional files. Administrators are still able to manually compress files from the Windows client.

By default, the VNXe File Deduplication and Compression feature is disabled for a storage resource. In the "Disabled" state, the storage resource has no deduplicated files and the policy engine does not scan for files to deduplicate.

An "Enabled" state indicates that the VNXe File Deduplication and Compression processing is enabled for the storage resource. When a storage resource is in the "Enabled" state, it may contain deduplicated files, and the policy engine scans the source resource for more files to deduplicate on its next scheduled run.

The "Suspended" state means that the VNXe File Deduplication and Compression processing is suspended for the storage resource. When a storage resource is in the "Suspended" state, it may contain deduplicated files. However, the policy engine does not scan to look for any more files to deduplicate.

Administrators can switch between deduplication states at any time. If they switch the deduplication state from "Enabled" or "Suspended" to "Disabled", the system checks whether there is sufficient space available in the storage resource to complete this process (without filling the storage resource completely) before accepting the request to disable deduplication.

If the space is not sufficient, the system informs the administrator that additional space is required to complete the operation, and recommends the extension of the storage resource.

Users can initiate a deduplicate scan manually by selecting the "Force Rescan Now" button.

Deduplication Settings

In Unisphere, policy settings are configured at the storage resource level. Administrators can configure the following policy criteria in the Deduplication tab that appears on the Details page of the file-level storage resource. There are two policy criteria that are configurable by the administrator:

- **Excluded File Extensions** – If certain file extensions need to be excluded from the deduplication process, set up the file extension exclusion list before running deduplication for the first time. If you add exclusions after deduplication has been run on a storage resource, the new exclusions take effect from that point forward and do not affect files that were deduplicated in previous scans. File extensions must include the leading dot.

- **Excluded Paths** – If you want to exclude pathnames from the deduplication process, add the pathname exclusion list before running deduplication for the first time. If exclusions are added after deduplication has been initiated on a storage resource, these exclusions take effect from that point forward. Files that were deduplicated in previous scans are not affected. The forward slash (/) is a valid directory delimiter within a single pathname.

Deduplication Status and Statistics

Figure 10 shows the results of the deduplication process on the data in the storage resource with the following statistics:

- **Status** – Running, In Progress, or Idle.

- **Last Scan** - Last successful scan of the storage resource was completed.

- **Files Deduplicated** – Number of files processed by the deduplication policy engine to save space and percentage of deduplicated files.

- **Complete** – Percentage of deduplication completed.

- **Total Size** – Total size of the storage resource.

- **Original Used Size** – Size of the storage resource before deduplication.

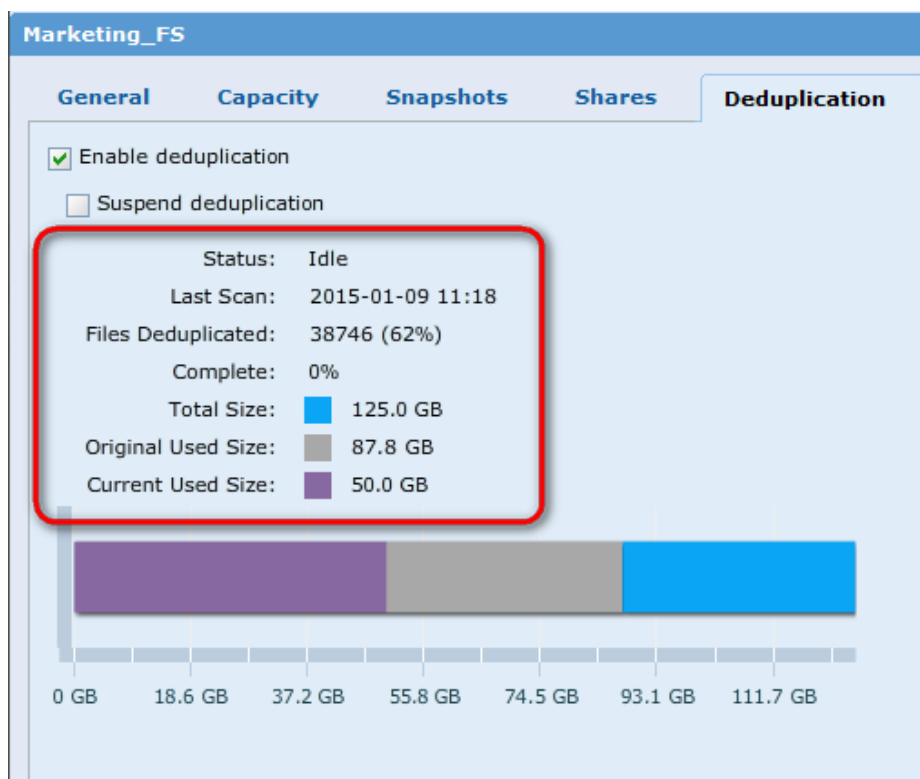- **Current Used Size** – Size of the storage resource after deduplication.

Figure 10: Status of Deduplication Scan

When scanning a storage resource enabled for deduplication for the first time, the statistics shown are reported in real time. After the first scan, statistics are reported as static values based on the last successful scan.

Client input/output to space-reduced files

The VNXe File Deduplication and Compression feature does not affect the client input/output (I/O) to files that have not been deduplicated. The feature does not introduce any additional overhead for access to files that it has not processed. The default policy is designed to filter out files that have frequent I/O access and thus avoid adversely affecting the time required to access those files.

Read access

Read access to deduplicated files is satisfied by decompressing the data in memory and passing it back to the client. The VNXe3200 storage system does not decompress or alter any data on disk in response to client-read activity. In addition, random reads require decompression of the requested portion of the file data and not of the entire file data. Reading a file that is deduplicated can take longer than reading a non-deduplicated file, because of the decompression activity. However, the opposite may also be true. Reading a deduplicated file is sometimes faster because less data needs to be read from the disk, which more than offsets the increased CPU activity associated with decompressing the data.

Write access

A client request to write to, or modify, a deduplicated file causes the requested portion of the file to reduplicate in the storage resource while preserving the deduplicated data for the remaining references to this file. The following three factors mitigate this effect:

- Most applications do not modify files. They typically make a local copy, modify it, and when finished, write the entire new file back to the file server, discarding the old copy in the process. Therefore, reduplication on the file server does not occur; it is just replaced.

- The VNXe3200 storage system avoids processing active files (accessed or modified recently) based on policy definitions. Hence, deduplicated files are less likely to be modified and, if they are, performance is less likely to be a critical factor.

- When a client writes to a deduplicated file, the SP writes only the individual blocks of text that have changed. The entire file is not decompressed and reduplicated on disk until the number of individual changed blocks plus the number of blocks in the corresponding deduplicated file is smaller than the logical file size. Therefore, the VNXe3200 storage system optimizes totally used disk space.

Interoperability with other VNXe3200 features

NDMP PAX Backup

When backing up over the network through CIFS or NFS, space-reduced files are filtered for their compressibility. This determines if the space-reduced files are to be backed up in their compressed format or in their original form. This is done so as not to affect the restore time on files that provide minimal amount of savings. Space-reduced files are neither reduplicated to their original size for transfer to the backup application nor are they reduplicated on disk. Although, the space saving storage efficiency benefits realized in the VNXe3200 storage system's deduplicated production storage do not flow through to backups when using network-based backups. When backing up deduplication-enabled storage with NDMP PAX, the compressed version of the file is backed up and all instances of the deduplicated file appear in the backup stream. It is worth noting that in most cases the majority of the space savings come from compression, which means that the majority of the space savings made in storage resource apply to the backup as well.

In NDMP PAX backups, a threshold is used to check on the compressibility of each space-reduced file before sending the file data compressed. There are cases when it is better to back up the file uncompressed, especially when there is sparsely written data associated with the file. By default, the deduplicated size of the file must be 90 percent (or less) than the original size of the file at the time it was deduplicated. This ensures that compressible files are backed up in their compressed format while non-compressible files are backed up in their non-deduplicated format so as not to sacrifice the restore time on files that would provide minimal amount of tape space savings.

When restoring files from a PAX-based NDMP backup, the files are restored in their compressed format. This feature also provides right-sizing capabilities in that an administrator can back up the deduplicated files from one storage resource and restore them onto a smaller storage resource. This feature is independent of the backup applications.

VNXe File-Level Retention

You can enable VNXe File Deduplication and Compression on enterprise File-Level Retention (FLR-E) storage without compromising on the protection offered to the data that the storage resources contain.

File Deduplication and Compression in a VMware Environment

File Deduplication and Compression can be used in a VMware environment to improve storage efficiency of the virtual machines (VMs). Active virtual machines are constantly changing states and are typically bypassed by the VNXe3200 storage system's deduplication policy engine.

VMs that are selected for storage optimization are compressed using the same compression algorithm that is used in VNXe File Deduplication and Compression. The

algorithm compresses the VMDK file (virtual disk) and leaves the other files that make up the VMs intact. When a compressed VMDK is read, only the data blocks containing the actual data are decompressed. Likewise, when data is written to the VMDK, instead of compressing it "on the fly", data is written to a set-aside buffer, which then gets compressed as a background task. In most situations, the amount of data read or written is a small percentage of the total amount of data stored.

FAST Cache and FAST VP

File deduplication and compression is fully compatible with the FAST Cache and FAST VP features. However, it is unlikely that deduplicated or compressed data will ever be in FAST Cache. This is because of policy criteria stating that a file must not be accessed for a certain period of time before it is eligible to be deduplicated and compressed. Due to this infrequent access, deduplicated and compressed data is likely to be relocated to lower tiers by FAST VP.

Deployment considerations

Consider the following when enabling VNXe File Deduplication and Compression:

- VNXe File Deduplication and Compression is supported with the UFS32 format only and cannot be used with VMware NFS Datastores created with the UFS64 format.

- You can enable VNXe File Deduplication and Compression on existing file systems and VMware NFS datastores (with UFS32 format). This feature cannot be applied to LUNs or VMware VMFS datastores hosted on the VNXe3200 storage system.

- VNXe File Deduplication and Compression does not deduplicate data across or between storage resources.

- With NDMP PAX, backing up a deduplication-enabled storage resource results in a shorter time window. However, the restore window required will increase.

- VNXe File Deduplication and Compression does not process or affect alternate data streams (also known as named attributes) associated with files and directories in the storage resource.

- VNXe File Deduplication and Compression does not process files less than 24 KB in size. The overhead associated with processing such files negates any space savings achieved.

- There is no limit to the size of a storage resource on which VNXe File Deduplication and Compression can be enabled. However, the storage resource must have at least 1 MB of free space before deduplication can be enabled. If there is not enough free space, an error message is generated.

- You can enable VNXe File Deduplication and Compression on any number of file-level storage resources.

- Since media files are not good candidates for compression, file systems containing a large amount of these files may not realize great savings. As a result, compressing these particular files on these file systems can lead to a decrease in system performance.

Conclusion

The VNXe File Deduplication and Compression feature adds to the VNXe3200 storage system's impressive storage efficiency capabilities by intelligently reducing space usage and providing further storage efficiency. You can optimize the use of the existing storage environment for storage resource data through file-level deduplication and compression. The savings achieved while using VNXe Deduplication and Compression varies as the savings depend on the type of files being stored on UFS32 file systems or VMware NFS Datastores.

A combination of file-level deduplication and compression represents the best technique to provide maximum benefit for the resources consumed. This feature builds on the VNXe3200 storage system's ease of use by providing a single-click option in Unisphere to enable deduplication for each storage resource.

References

The following can be found on the EMC Online Support website:

- Introduction to the EMC VNXe3200 – A Detailed Review
- EMC Unisphere for the VNXe3200: Next-Generation Storage Management – A Detailed Review