WHITE PAPER: DATA AND SYSTEM PROTECTION

Symantec Enterprise Vault™

File System Archiving

Daniel Maiworm

May 1, *2007*

White Paper: Data and System Protection

Symantec Enterprise Vault

File System Archiving

Contents

Introduction

This white paper is designed to give technical staff, project managers, and Microsoft® Windows® file server experts an overview of Symantec™ Enterprise Vault File System Archiving (FSA). It is not intended to introduce File System Archiving to a nontechnical audience.

Overview of File System Archiving

More than hierarchical storage management

For more than 25 years, organizations have been attracted to the idea of providing different storage layers according to the relevance of data, an approach referred to as hierarchical storage management (HSM). Today, however, companies accumulate an enormous amount of information, presenting new challenges. In addition to simply storing the information, they must search and retain the right type of information according to defined policies.

Enterprise Vault has been designed as a universal repository and management solution primarily for unstructured data. It was started as a project within Digital Equipment Corporation (DEC) in 1997. The first version shipped in 1999, and since then Enterprise Vault has won numerous awards and analyst recommendations. Even more important, over 5,000 global customers use Enterprise Vault as an archiving framework to reduce risk and increase operational efficiency.

Enterprise Vault exceeds the limitations of traditional HSM applications by providing facilities to index and classify data, discover it for legal purposes, and facilitate compliance with external regulations. To provide a complete life-cycle management solution, integration of file blocking and state-of-the art quota management for file servers is on the near-term roadmap.

While many competing archiving solutions consist of a portfolio of point solutions acquired or delivered by OEMs that are more or less integrated on the surface, Enterprise Vault has been developed as a flexible framework using a single storage engine, unified administration console, and a security subsystem that matches the flexibility of existing e-mail and file servers (see Figure 1).

Today the majority of systems that have been developed for archiving business records struggle with the unstructured nature of e-mail and files. Offline access, permission synchronization, object-level retention, and full-text indexing are often afterthoughts. In contrast, Enterprise Vault has been developed with these requirements in mind, delivering enterprise-class performance with minimal impact on end users.

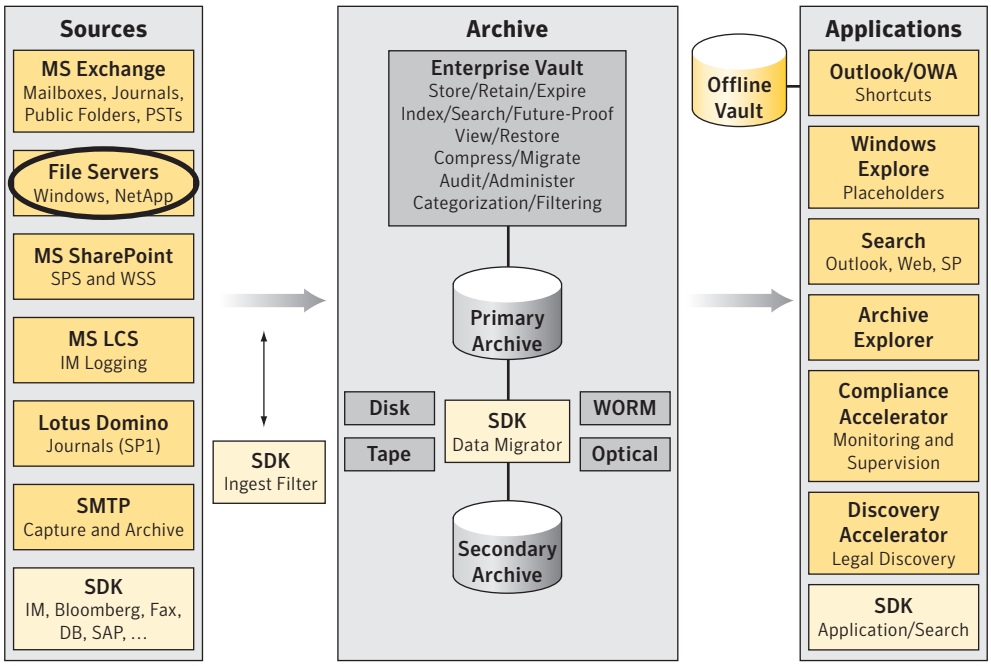

Figure 1. Enterprise Vault archiving framework

Differentiators

Enterprise Vault is a unique product in the marketplace owing to features such as:

- Single point of administration for all archiving targets
- Unified storage system for email messages, SharePoint® files, and other data
- Flexible rules engine
- Integrated Active Directory security/authentication system
- Comprehensive recall limits, preventing abuse
- No application logic on the file server (retrieval agent only)
- Advanced monitoring and reporting
- Data compression
- File versioning and pruning
- Single-instance storage, independent of the back-end platform

Optional features (provided by FSA Archiving and Search option):

- Full-text indexing and search
- Alternate access through Archive Explorer (independent from placeholders)

- HTML renditions for long-term information access
- Legal discovery option through Enterprise Vault Discovery Accelerator (licensed separately)

In summary, Enterprise Vault offers the functionality of a comprehensive information management platform, rather than a classic HSM solution that focuses on data and storage management alone.

High-level concepts

Administration

Enterprise Vault administration console

File System Archiving is a core part of the Enterprise Vault framework and is administered from the same Microsoft Management Console (MMC) snap-in as the other archiving targets (see Figure 2). The use of MMC technology allows administrators to add the Enterprise Vault administration interface to their existing set of console snap-ins, reducing the complexity and enhancing the productivity of daily operations.

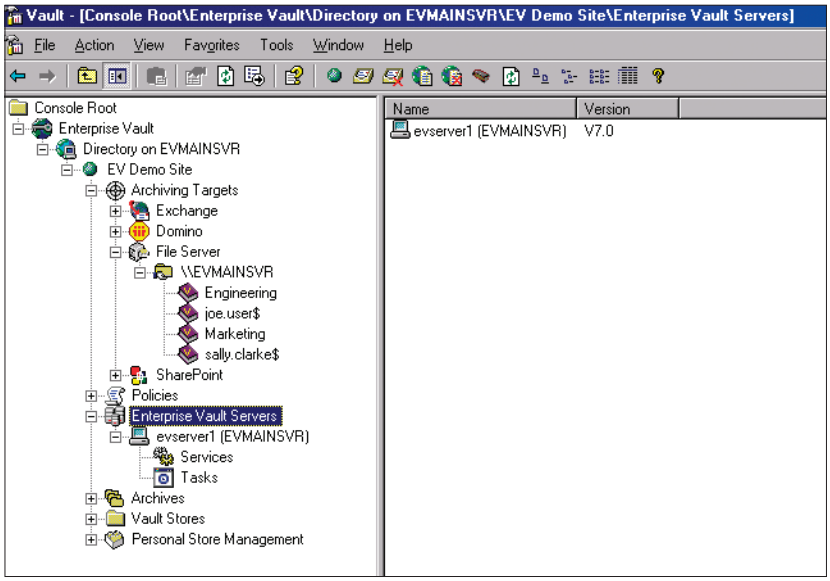

Figure 2. MMC admin console (power admin role)

It is possible to restrict access to functionality in the Enterprise Vault admin console based on the role of the administrative user. In fact, you can show a file server admin only the containers and options that are relevant to his or her job, while hiding features like Exchange and Domino® to prevent accidental changes to policies outside the administrator's defined scope (see Figure 3).

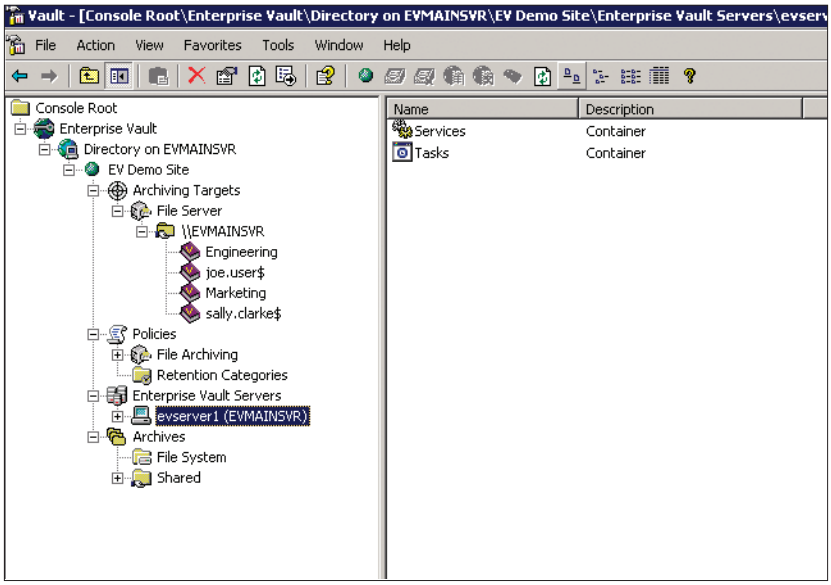

Figure 3. MMC admin console for the file admin role

In addition to the administration console, there are a number of command-line utilities to automate the provisioning of archive points and start an archive run from a script. For example, an archive run can be triggered once the file server backup finishes.

Another powerful command-line application is FSAUtility.exe. This tool helps with the consolidation of Windows file servers by moving placeholders without recalling the original files. It also provides functionality to re-create corrupt or missing placeholders (for example, after a backup restore). Specifically, FSAUtility.exe can:

- Re-create missing placeholders (checks consistency between file server and archive)
- Re-create (export) original files
- Re-create lost/corrupt archive points according to the archive configuration
- Move placeholders without triggering file recalls (for moving file shares or consolidating servers)

The exact syntax for FSAUtility.exe can be found in the Enterprise Vault online help. It is also briefly described in Appendix B. Note that FSAUtility currently supports only Windows file servers. Support for running FSAUtility with a NetApp® filer or EMC Celerra® device is planned for addition in a future release.

Reporting and monitoring

Enterprise Vault is tightly integrated with the Windows Event Viewer and therefore can be used with all applications that analyze and monitor the event logs. It adds a new log to the Windows Event Viewer to provide full status information without overloading the existing application log.

With Enterprise Vault Operations Manager, administrators can monitor the status and health of all Enterprise Vault servers within a site from a central Web-based operations console. This provides instant information about the availability of Enterprise Vault and helps administrators meet even the highest service level agreement (SLA) requirements.

For organizations that want to use an external monitoring framework, Symantec provides out-of-the-box integration with Microsoft Operations Manager (MOM), enabling you to monitor the status of Enterprise Vault with the same tools and systems that you use for monitoring the availability of Windows file servers. Other management frameworks can either be modified to monitor the Enterprise Vault event logs and service state or integrated to use MOM as a management agent for Microsoft servers and Enterprise Vault.

To convey the effectiveness of Enterprise Vault and to justify the investment to the business, Enterprise Vault incorporates comprehensive reporting based on Microsoft SQL Server Reporting Services, delivering instant reports (for example, on monthly archived volumes) with options for custom reports and designs.

In addition to SQL Server–based reporting, comprehensive report files can be created either during the production run or in a report mode, which produces a "what if" analysis without changing the information on the file server. This integrated reporting facility allows administrators to quickly identify which files get archived by which policy and rule, what volume of information can be migrated from the file server, and whether the files have any explicit permissions.

Information access

Placeholders (Windows, NetApp, and EMC Celerra)

With Windows based file servers, Enterprise Vault provides completely transparent access for end users by using placeholder technology, an extension of the NTFS offline mechanism used by the Windows Remote Storage Service (RSS). As shown in Figure 4, placeholders display the same icon as the original file, with a small clock as an overlay symbol to indicate that the file has been archived and that access to the information might take slightly longer.

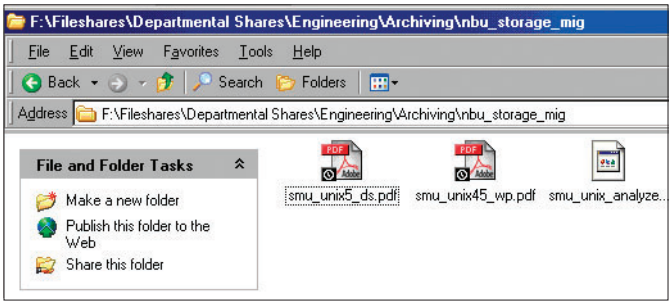

Figure 4. Enterprise Vault placeholders (Windows, NetApp, and Celerra)

A registry key on the file server lets you choose whether you want to display the original file size (the size before archiving) or the placeholder's true size of 0 bytes. The default option is to display the original file size (see Appendix A for details).

It is possible to restore corrupt or missing Windows placeholders from the archive and to check the consistency of the placeholders by using FSAUtility.exe (see Appendix B for details).

To help with the deployment of Enterprise Vault—especially in distributed environments—it is possible to "push" install the placeholder service on some or all file servers, without the need to log in locally on each file server machine. This streamlines deployments in larger organizations and reduces the burden of managing large numbers of file servers.

Placeholders can also be used on network attached storage (NAS) from Network Appliance and EMC. With version 7.0 of the ONTAP operating system, Network Appliance has integrated the Enterprise Vault placeholder technology into its NAS family. This allows end users to have the same transparent access to files archived by Enterprise Vault as they enjoy for documents still stored on the filer itself (see Figure 5).

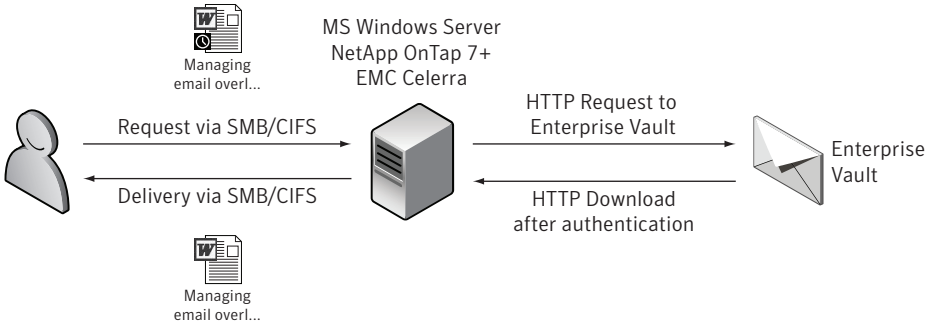

Figure 5. Placeholder recall mechanism for Windows , NetApp, and EMC Celerra

Before adding a NetApp filer, you must give the Enterprise Vault service account administrative permissions on the NetApp filer. After you set the permissions, you can add the NetApp filer as a file server from the Enterprise Vault admin console. If you have a NetApp filer with Data ONTAP 7.0, you can archive files and leave Enterprise Vault placeholders. If placeholders are supported on the NetApp filer, the configuration wizard will give you the option of enabling them; otherwise, this option will be dimmed.

In addition, Symantec has implemented runaway recall limits for archiving from Network Appliance filers to ensure that a single client cannot recall a large number of files in a short period of time. This avoids excessive recalls due to antivirus scans, backups, and so on.

With version 7.0, Enterprise Vault has added support for archiving from the EMC Celerra platform, similar to the integration described for Network Appliance devices. To provide the same seamless access to data archived from a Celerra NAS device, Enterprise Vault delivers comprehensive integration with the Celerra Filemover (DHSM) API, including the capability of tracking delete actions (delete-on-delete) and optional pruning when multiple versions of an item have been archived. This integration is available on all Celerra systems that run the DART NAS operating system in version 5.5 or higher.

Note that the ExcludedExe functionality does not exist for NAS devices, and all archived file placeholders will display a file size of 0 bytes to end users.

Internet links

When archiving files from NAS systems that do not support Enterprise Vault placeholders (note that the NAS still needs to be 100 percent CIFS compatible to be archived), Enterprise Vault can still provide direct access to the archived files by placing a download link (*.url file used by Microsoft Internet Explorer) to replace the original file after archiving. For example, a document called MyDocument.doc will be visible as MyDocument.url, with the icon changed from a Word document to the Internet Explorer symbol, as shown in Figure 6.

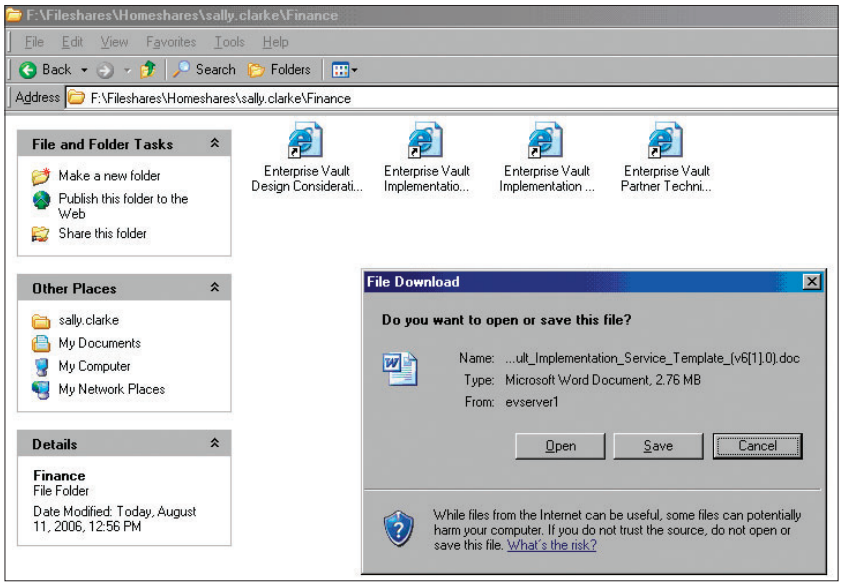

Figure 6. Enterprise Vault Internet links

When you double-click the Internet shortcut, the client's Web browser provides the option to either save or view the item from Enterprise Vault. Although this does not provide the same transparency as the placeholders, it is a valid way of providing access, especially if the documents are old and rarely accessed (see Figure 7).

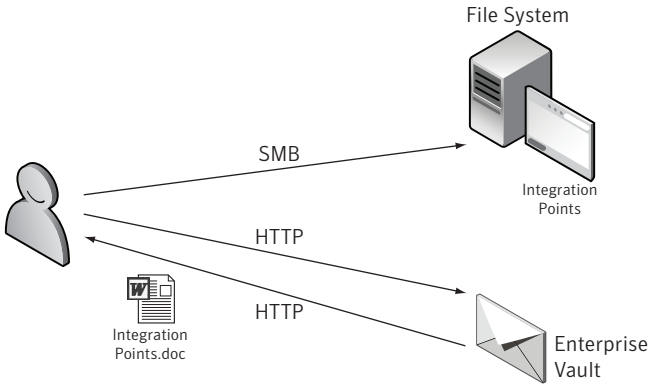

Figure 7. Enterprise Vault Internet links end-user access

This mechanism is also often used as a simple interface for third-party integrations and allows custom solutions to extract URL information from the file and store those Enterprise Vault download instructions directly within the custom application.

Archive Explorer

A completely separate way of accessing items stored within Enterprise Vault is the Archive Explorer Web client. This allows administrators to completely remove old documents from the file server and provide access to this information via a Web browser. As shown in Figure 8, the end user can browse the hierarchical folder structure based on the names of the folders from which files were archived. This provides the ability to completely remove the dependency between the file server and the archive and is a valid method to clean up information that has been archived for years.

As Archive Explorer also displays the user's archived data from Exchange mailboxes and public folders, it can serve as the "one-stop" client for your company's historic information.

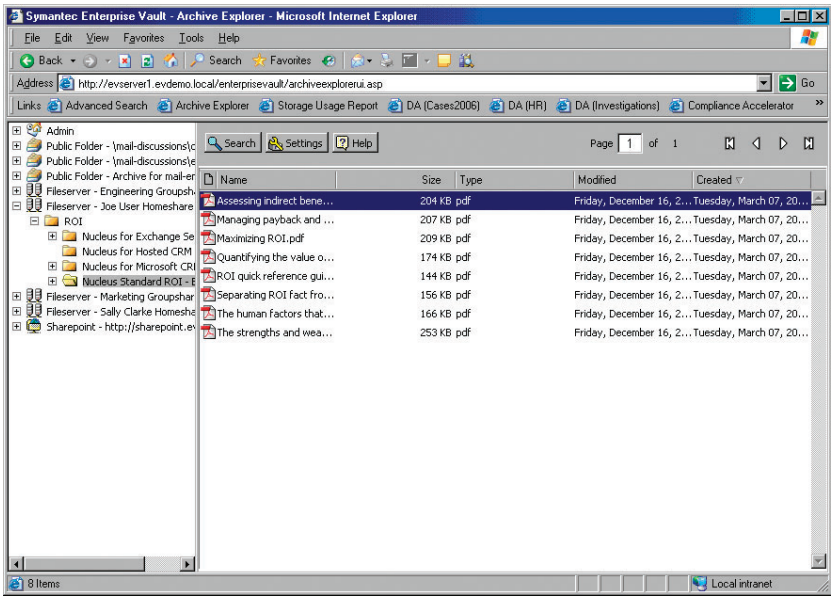

Figure 8. Enterprise Vault Archive Explorer

Storage

Supported storage platforms and tiered-storage solutions

For the back-end storage of archived content, Enterprise Vault supports almost all major storage technologies on the market. This includes any NTFS/CIFS disk-based system like DAS, SAN, and NAS, as well as other storage solutions including CAS, tape, DVD, and optical. Special integrations with write-once/read-many (WORM) storage systems provide an unparalleled choice of solutions from the leading storage vendors, including:

- EMC Centera
- EMC Celerra NSX
- Fujitsu ETERNUS®
- Hitachi® Content Archive Platform (HCAP)
- IBM® DR550
- IBM Tivoli® Storage Manager (TSM) with Data Retention Manager
- NetApp SnapLock® (on NetApp filer and NearStore)
- Pillar Axiom®
- Sun StorEdge™ 5310

You can potentially use all supported storage systems in parallel to provide different storage features for different types of target data.

Since Enterprise Vault was developed as a single, integrated archiving framework, all data sources such as mail, files, and portal data can be archived to the same storage systems and locations, which are controlled by the same granular migration and retention policies and administered from the same MMC console.

Note that Enterprise Vault has been designed to integrate with both disk and tape technology to provide a balance between accessibility of information and cost of storage. Therefore, it is possible to move historic information that is rarely requested to a secondary storage system as well as to collect smaller items into container files for more efficient storage and easier management and backup of the stored data. The Enterprise Vault secondary storage option can now be integrated with Veritas NetBackup™, IBM DR550, and Fujitsu ETERNUS WORM storage, as well as an open interface to move the data to another file system that is based on less expensive disks or optical media.

Before items can be moved to the secondary storage, they must be collected into CAB files to optimize the storage footprint and to limit the number of objects created in the media server catalog. As the CAB file needs to be recalled to the primary storage location, direct read and direct write to tape are not supported. For performance reasons, it is recommended that you only use tape for data that is seldom or never accessed.

Note that you cannot collect archived data stored on an EMC Centera device or migrate it to another storage device.

Compression and single-instance storage

Enterprise Vault creates a unique MD5 fingerprint for every file that is archived. If multiple files have the same hash code, only one copy of the file is physically stored. Many organizations see a dramatic reduction in overall storage volume, as hundreds of archived documents can often be consolidated into a single instance in Enterprise Vault.

Before storing any files, Enterprise Vault compresses the information with the ZLib compression standard to optimize storage efficiency. This often leads to a 30 percent to 40 percent size reduction when archiving a standard mix of office documents. As Enterprise Vault performs compression and full-text indexing on files using the server's RAM, there is a configurable limit on the size of files to which this kind of additional processing can be applied.

Files larger than this configured limit will still be archived, but only the file name and the metadata will be indexed, and no compression will be performed. You can easily identify archived files larger than this threshold because they are stored in the archive as a DVF instead of a DVS file. The default threshold for large items is 50 MB, and this should not be changed.

Versioning and pruning

Enterprise Vault provides file versioning and pruning out of the box. This is especially useful when changes to a file need to be preserved over the lifetime of the document. Previous versions can be accessed via the Archive Explorer Web interface as well as the various Enterprise Vault search interfaces.

In Enterprise Vault FSA, it is easy to limit the number of versions kept in the archive by setting a maximum number of versions in the FSA task properties, found in the administration console. This removes the oldest versions until the preset limit is reached when the FSAPruningTask runs according to schedule.

The option to run pruning via the EVFSARunNow utility from the command line is also available for script and batch purposes (see Appendix B).

Indexing

Indexing benefits

Today, search technology is one of IT's hottest topics and greatest challenges. The Internet has shown that storing information is only one part of the solution, especially considering the size and availability of today's storage systems.

The challenge is to enable users to retrieve the right piece of information out of the ever-growing haystack of information, while maintaining security and privacy. Given the Enterprise Vault design heritage, it was an obvious choice to integrate the widely known and enterprise-class AltaVista search engine as a core component of the archiving framework. Enterprise Vault FSA not only uses AltaVista, it also proactively manages the complete index subsystem. This means that Enterprise Vault has additional intelligence to split the data into meaningful subparts, roll over large indexes into new index volumes, and reindex only a single small index if a corruption occurs.

Additionally, Enterprise Vault validates the data before indexing it to prevent meaningless binary data from being added to the AltaVista indexes. This further reduces the index storage footprint in comparison with competing products. For more information about AltaVista, refer to the technical white paper on Enterprise Vault indexing, available on symantec.com.

Archive points

Archive points represent a tag within the file system that marks the beginning of a new index subtree. For example, a user home folder represents a logical collection of files that are probably only relevant to the searches of the particular owner. Therefore, marking the root folder of this subtree as an archive point will instruct FSA to create a separate archive for this data, keeping the index small and fast while providing a separate search target for the end user.

Archive points can be conveniently set and managed from the Enterprise Vault administration console, including the option to auto-enable archive points for subfolders. It is also possible to completely disable indexing, per archive point. This provides greater archiving performance, but because archived files are not indexed, you cannot search for items that have been archived or use Archive Explorer to view them.

As stated previously, full-text indexing is only available for files that can be processed in server RAM. For files larger than a configured threshold, only the metadata and file name are indexed.

For scripting and batch processing, you can also use the ArchivePoints.exe command-line utility. (See Appendix B for more information about the ArchivePoints utility and how to disable indexing.)

Enterprise Vault business accelerators

Enterprise Vault Discovery Accelerator is fully compatible with File System Archiving. This provides not only full-text searching of all archived file content, but also export and/or production of the search results to a designated storage location, while preserving native file formats and folder structures. For more information about Discovery Accelerator, refer to the Symantec technical white paper titled *Reducing E-Discovery Cost and Risk with Discovery Accelerator*, available on symantec.com.

Note that Enterprise Vault Compliance Accelerator can search for keywords found in archived files, but it cannot take a random percentage sample of newly archived files as it does with Exchange journaling.

Data protection

Security

Service account requirements

All services and processes in Enterprise Vault run in the context of a domain user account that needs to be a local (machine) administrator on the Enterprise Vault server. This Enterprise Vault service account will be granted the "Log-On as a service" permission on the Enterprise Vault server to allow it to run as a service while no user is interactively logged in during archive runs.

Enterprise Vault archives files over the network. Therefore, full control permissions on the network share and file system permissions need to be set on the file server being archived. The Enterprise Vault service account needs to be able to change files over the CIFS/SMB protocol (share permissions), as well as create the placeholder or URL link (file system permissions). Note that during installation of the placeholder service, the setup routine automatically adds the Enterprise Vault service account to the local administrator's group on the file server.

As Enterprise Vault does not allow anonymous access to any information stored by the system, all clients that try to access objects stored by Enterprise Vault must authenticate themselves. Therefore, the placeholder service (which acts as an Enterprise Vault client on the file server on behalf of the requesting user) needs to pass its credentials to the Internet Information Server (IIS) that runs on the Enterprise Vault server. To do so, the Internet settings on each server running the Enterprise Vault placeholder service must store the name of the FSA server in the list of local intranet Web sites. The placeholder service Push-Install wizard normally configures this automatically, but it may be helpful to verify the presence of this server name during troubleshooting, as shown in Figure 9. Note that this Internet configuration is profile specific, so it must be done in the context of the Enterprise Vault service account.

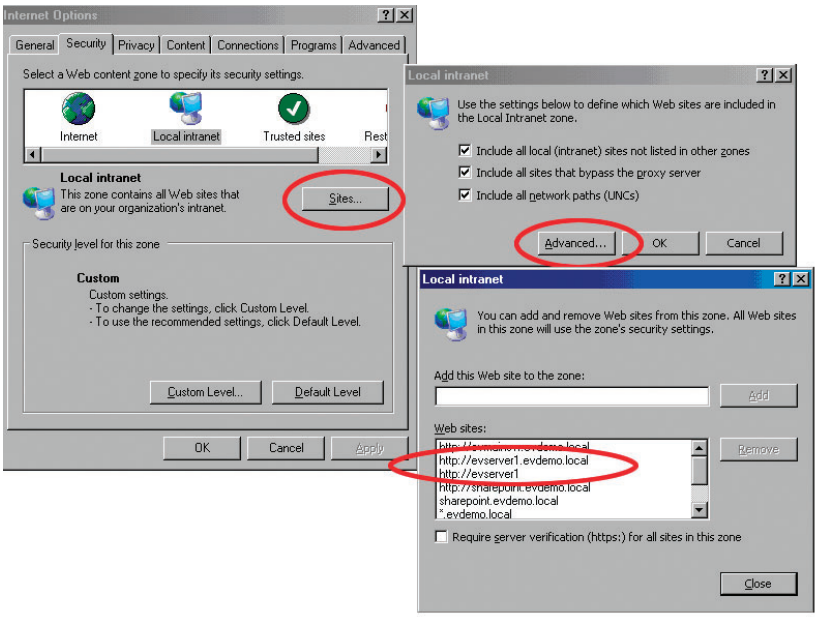

Figure 9. Internet Explorer configuration on each Windows file server for automatic HTTP authentication by the placeholder service.

Permissions stored in Enterprise Vault (folder-level permissions)

Enterprise Vault was created with the highest level of security in mind; therefore, permissions on the file server are synchronized twice per day, and end-user requests for archived files are authenticated accordingly. For performance reasons, Symantec has standardized on folder-level rather than file-level permissions. So, as files with explicit file-level permissions are encountered during archiving, Enterprise Vault can be configured (on a per-policy basis) to either ignore the file (do not archive it at all) or to archive the file with the access permissions of the parent folder that contains it. Companies are advised to think carefully about this configuration option before implementation. There may be certain folders with sensitive content, such as HR or accounting folders, that should use a policy that ignores files with explicit permissions. Other folders, such as end-user home directories, could use a policy that archives files with explicit permissions, applying the permissions of the parent folder.

Folder tracking via CIFS streams (EVFolderXML)

In order to track changes on the file server, such as folder reshuffling within the hierarchy, Enterprise Vault adds an invisible, unique stamp to every folder from which it archives. These stamps are tracked during the life cycle of the folder and get synchronized with the archived content. By using CIFS streams to store the unique folder ID, this works seamlessly within the same volume, as long as fully compatible NTFS file systems are used. Note that the path of the original folder will not be updated, so the search results and Archive Explorer structure will still reflect the old path.

Availability

Enterprise Vault and high availability: Clustering and failover

Enterprise Vault can be configured as a highly available solution to deliver continuous operation in the most critical environments though its support for market-leading Veritas™ Cluster Server and Veritas Storage Foundation™ high availability technologies, as well as Microsoft Cluster Server (MSCS) on Windows Server 2003 Enterprise Edition. This allows for configurations with virtually no downtime. Note that an Active/Active setup of cluster nodes is not supported in Veritas Cluster Server or MSCS configurations.

In addition to the high availability solution using cluster technology on the OS level, Enterprise Vault provides a manual failover on the application level without any additional software. A simple reconfiguration of the DNS server that manages the Enterprise Vault server's IP lookup from the placeholder service, combined with a failover procedure invoked with a simple mouse click from the Enterprise Vault admin console, will make the secondary system available in a few minutes. As most of the historic information in the archive does not have the same business value as the most recent documents, this option is acceptable for most enterprise environments. In addition, it can be used in Active/Active configurations so all Enterprise Vault servers can be used in normal operations without the need for idle failover machines.

Enterprise Vault has also been designed to support clustered file servers. This normally complex task incorporates an easy-to-use wizard within the user interface that guides administrators through the setup process. Enterprise Vault also supports the following cluster configurations on the file server: Active-Passive, Active-Active, and 3-node and 4-node clusters with and without a passive node.

The Enterprise Vault placeholder service can be installed on both Microsoft Cluster Server and Veritas Cluster Server, either as a managed or unmanaged resource. Various Symantec technical documents offer a detailed overview of wizard-guided configurations according to specific environments (see Appendix C).

Archiving from DFS shares

In version 7.0, Enterprise Vault has limited support for archiving from DFS shares. For the latest compatibility information about Enterprise Vault FSA and DFS, please consult the technical note published on the Symantec Support Web site.

Administration concepts

Policies

Enterprise Vault policies are normally set on a volume level. A volume is normally a Windows (CIFS/SMB) network share, as very few organizations would implement archiving from the Enterprise Vault server's local drives. Policies consist of quota policies, placeholder type and deletion strategy, retention category, archiving rules, and the handling of files that contain explicit permissions.

To allow for more granular policy configuration within a share, Enterprise Vault also has the ability to create policies on a folder level. These policies can either complement or completely override the volume policies that would normally control the folder's archiving behavior. Folder policies are often used to disable the automatic archiving of system data or sensitive data that should not be archived without the user's consent.

When storage savings is the primary goal and volume quotas are currently configured in the file system, Enterprise Vault can use a policy that starts archiving only if a high-water mark of available disk space on the target partition of the file server is reached. For example, if the file server has data stored on a 200 GB NTFS volume, you can set a high-water mark of 90 percent that will start archiving only if less than 20 GB of space is available on the volume. It is also possible to specify a low-water mark that controls when archiving will stop again. For example, a low-water mark of 60 percent would cause Enterprise Vault to continue archiving additional files until at least 80 GB of disk space was available on the 200 GB volume. After that, it would monitor the free space until the high-water mark was reached, and then the process would begin again.

The ability to control archiving by quotas set on group and user levels is on the roadmap for a future version of File System Archiving.

File groups

Rather than listing file extensions in every policy that is created, you can define one or more reusable file groups to list common file extensions that will be specified in multiple policies. A typical example would be to create a group called “Office Files” and add the extensions *.doc, *.xls, and *.ppt. These groups can be referenced later during policy creation, simplifying rule management and increasing consistency and accuracy over time, as new file types must be added to all archiving rules. When a new file type is added to Microsoft Office, you can simply update the file group instead of changing dozens of policies. Figure 10 illustrates a list of preconfigured file groups.

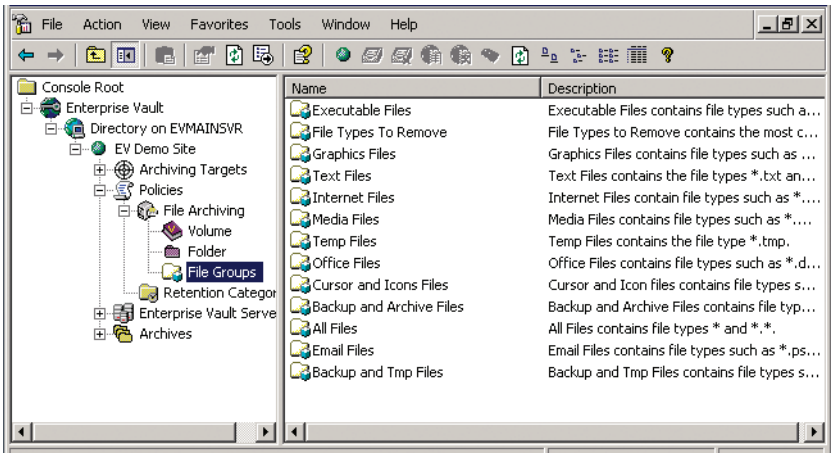

Figure 10. Default file groups shipping with Enterprise Vault

Rules

Within a policy, you can specify multiple rules that target certain files for an archiving action. The possible actions are archive, do not archive, delete, and archive copy and reset, which resets the standard NTFS file archive bit. The first three options are self-explanatory, while the last provides an option to create archived copies of files that have changed. Note that this might interfere with your backup strategy if the file server's backup application relies on the status of the NTFS archive bit. Figure 11 shows a policy with several different types of rules configured.

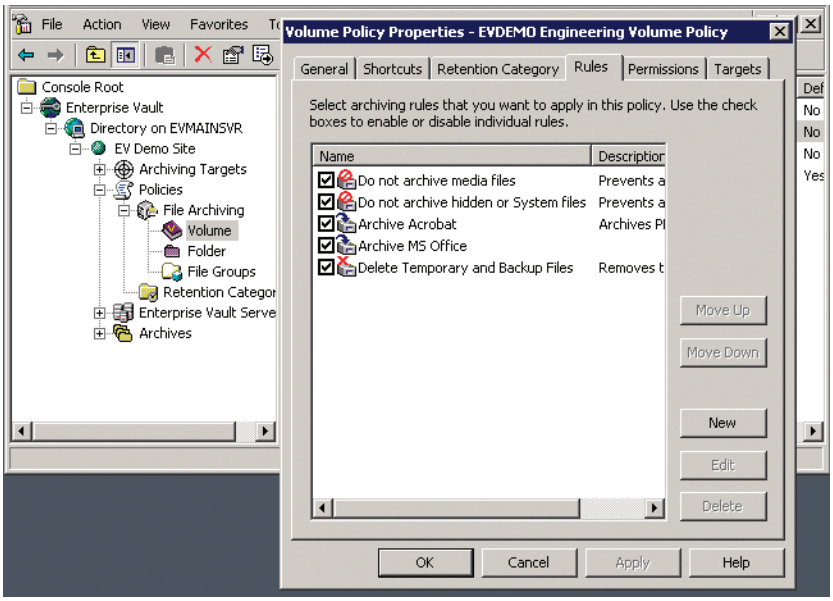

Figure 11. Rules view within a policy

When creating rules, you can set the criteria based on file name (or extension), size, time, and other file attributes and properties. To specify the files that should be included in the rule, you can either refer to an existing file group or manually enter the rule-specific file names and extensions. In both the file group and rule properties, the file name filter can include wildcards and multiple entries separated by a comma (for example, *.doc,*.xls,*.ppt).

The file size can be configured as greater than/smaller than in kilobytes, while the time filter includes “Last accessed,” “Last modified,” and “Created” NTFS stamps. Finally, the Attributes tab allows you to select various NTFS attributes such as hidden, system, compressed, and read-only.

The rules are processed according to their priority in the list, from top to bottom. Thus Enterprise Vault compares each file with the archive settings in the top rule before moving on to the next rule if there was no match. Once a criteria match is found, subsequent rules are not considered. This approach allows maximum flexibility in creating the archiving policy while minimizing administrative burden.

Targets

Microsoft Windows file servers

Enterprise Vault supports target file servers running Windows 2000 Server and Windows Server 2003 operating systems. This includes archiving from systems based on the Windows Storage Server 2003 NAS operating system, but note that your NAS vendor may hesitate to support systems that run the Enterprise Vault placeholder service on their appliance even though it is running Windows Storage Server.

On file servers that run Windows, the Enterprise Vault placeholder service controls the recall of archived items. This service runs in the security context of the Enterprise Vault service account and is designed to transfer requested files back to the file server by using the HTTP protocol.

You can enable shares on the target explicitly by specifying the UNC path or by using an auto-enabler, which can be created at the user level to automatically enable any new folders created underneath. This is intended for use with users’ home folders (\\share\folder\users\homeshare...).

NetApp filer (ONTAP 7.0)

Enterprise Vault supports archiving content from Network Appliance filers that run the NetApp ONTAP 7.0 (or higher) operating system. This also includes NetApp NearStore systems.

In order to add a NetApp file server to Enterprise Vault for archiving, first you need to give the Enterprise Vault service account administrative rights on the filer. If configured correctly, Enterprise Vault will create an FPolicy on the filer that orders it to trigger a file restore via HTTP whenever an offline file is requested.

Note: Automatically configuring “virtual filers” is only supported in ONTAP 7.2 and higher.

Some Windows functionality like the ExcludedExe mechanism is not available when archiving from NetApp file servers because NetApp does not run third-party applications directly from its systems. Also, the archived files always display a file size of 0 bytes.

EMC Celerra NAS (DART 5.5 plus Filemover API)

Enterprise Vault fully supports archiving from EMC Celerra NAS systems that run the DART 5.5 operating system. If you intend to use placeholder shortcuts on the Celerra device, you must enable the FileMover functionality on Celerra. To configure the device, log on to the Celerra Control Station and add an account for Enterprise Vault to use for authentication on Celerra. After you have enabled a particular file system for the Celerra FileMover, configure the Data Mover to accept Celerra FileMover API connections and specify the HTTP details used for recalls from Enterprise Vault.

Note: For tracking deletions, Celerra Logging also needs to be enabled and configured.

After completing the preparations, you can use the administration console to add the Celerra device as another archiving target.

Tasks

The archiving task is the component that processes the file server's content according to a schedule and a set of policies. The archiving task for FSA is a subprocess of the task controller and does not rely on a one-to-one relationship between file servers and tasks, which provides enhanced scalability and flexibility. Multiple FSA archiving tasks can archive from a single file server; conversely, a single FSA archiving task can archive from a number of different file servers.

Each archiving task has its own schedule. In Enterprise Vault, you can specify the time when either a file server or a share on the server should be processed. This is done via the standard Microsoft control. The scheduling will start at the selected point in time and end when the scheduled end time is reached. Note that File System Archiving will not pick up where it stopped on the last run. It will rescan the volume or share again during the next archive run.

In the task properties, the administrator can configure the task to run only in report mode, along with the maximum number of report files to retain. In addition to the report files on disk, the task properties have a status page that provides information about the progress of the current archiving run (for example, the number of files and folders processed so far).

FSA and third-party products

Antivirus and backup

FSA has been certified to work with many leading backup and antivirus products. All enterprise-class backup products and virus scanners for Windows file servers (such as Veritas NetBackup and Symantec AntiVirus™) should have the option to properly handle offline files without triggering a recall for each placeholder that is touched. If a backup or antivirus application has known issues either with the Microsoft Windows Remote Storage Service or the NTFS offline functionality, contact the vendor for an updated version.

If no updated version is available, you can use the ExcludeExe registry key on the file server to prevent the Enterprise Vault placeholder service from serving any requests issued by the application that might cause problems with file recalls.

Quota managers and storage resource management solutions

The Enterprise Vault File System Archiving component is compatible with the built-in quota management of Windows Server 2003 as well as the Veritas Storage Exec™ quota management facility.

As the APIs for quota management are not so clearly defined, Symantec cannot guarantee full compatibility with other third-party solutions. For more information about software compatibility, please refer to the Enterprise Vault certification tables available on http://support.symantec.com.

FSA may not be compatible with tools that calculate and report file sizes, such as many of the common storage resource management (SRM) solutions. This is because FSA uses a placeholder that is less than 4 KB, and it emulates the file's original size as it was reported prior to archiving. Therefore, only tight integration of Enterprise Vault FSA with Storage Exec reporting capabilities is certified to produce accurate results when calculating the size of an archived file system.

Mechanisms to prevent unwanted recalls (Windows only)

If files are recalled via a network connection, there is no way to determine whether the recall was initiated by an end user or an application. Most file-scanning applications (such as backup or antivirus products) recognize the offline bit on placeholders and do not recall the archived item when accessing the placeholder object. However, since some applications do not honor the offline bit properly, Enterprise Vault provides application-level protection against unwanted recalls. The default setting prevents more than 20 file requests in a 10-second interval. See Appendix A for details on how to configure recall limits.

For Windows based file servers, you can also specify a list of programs that are prohibited from recalling archived items. This would most likely be useful if you use an antivirus or SRM program that does not honor the offline file attribute. Note that this functionality will only work against applications that are stored and executed locally on the file server (see Appendix A for details).

Another way to block the mass recall of files for scheduled events like nightly backups would be to use a command-line utility called EVFSABackupmode.exe on the file server before and after running the process, such as backup, that could potentially cause a recall problem. When the file server is in backup mode, it does not recall any files from Enterprise Vault. Note that backup mode does not affect archiving; it merely stops all files from being recalled. If the account that recalls the files can be identified, you can create an Enterprise Vault Backup Operators security group that gets blocked, while normal users can still recall files, even in backup mode.

Conclusion

Symantec Enterprise Vault File System Archiving is clearly unrivaled in the Windows file system archiving market, thanks to well-designed features such as granular policy management, single-instance storage, seamless end-user access, full-text indexing, Web-based visibility into an archived file system, protections for mass file recalls, and more. Enterprise Vault surpasses the performance of typical file management applications, which merely move data based on broad policies from one storage device to another. Plus, as files are now routinely being requested and searched for internal and external investigations, Enterprise Vault enables organizations to include file system content as a key component of their enterprisewide archiving and discovery solution.

Appendix A: Registry keys

Emulating the original file size

Key name:
FileSizeEmulation

Location:
HKEY_LOCAL_MACHINE\SOFTWARE\KVS\Enterprise Vault\FSA\PlaceholderService

Content:
DWORD

Possible values:
0—Placeholder files show a size of zero.
1—(Default) Placeholder files show the size of the original files.

Description:
Controls whether placeholders appear to have a size of zero or sizes that match those of the corresponding archived items.

Note: This key only works on Windows file servers.

Excluding local applications from recalling placeholders

Key name:
ExcludedExes

Location:
HKEY_LOCAL_MACHINE\SOFTWARE\KVS\Enterprise Vault\FSA\PlaceholderService

Content:
String

Description:
For file servers other than NetApp filer devices, it is possible to specify a list of programs that are prohibited from recalling archived items. This is most likely to be useful if you use an antivirus or backup program that does not honor the file system offline attribute.

To specify a list of prohibited programs, edit ExcludedExes to specify the names of the program executable files, separated by semicolons (;). For example, to exclude Windows Explorer, MyBackupProg.exe, and a program called Antivirus.exe, you could specify:

Explorer.exe;MyBackupProg.exe;Antivirus.exe

If you change the list of prohibited programs, you must restart the placeholder service to effect the change. It is possible to specify a list of programs that are prohibited from recalling archived items.

Limiting the number of recalls per time interval

Key name:
RecallLimitTimeInterval

Location:
HKEY_LOCAL_MACHINE\SOFTWARE\KVS\Enterprise Vault\FSA\PlaceholderService

Content:
DWORD

(Default value: 10 seconds)

Description:
RecallLimitTimeInterval specifies the number of seconds in which a maximum of RecallLimitMaxRecalls recalls is allowed. When this limit is reached, there is an additional wait of RecallLimitTimeInterval seconds before the count is reset.

Note: This feature is not available when archiving from Network Appliance NAS devices.

Key name:
RecallLimitMaxRecalls

Location:
HKEY_LOCAL_MACHINE\SOFTWARE\KVS\Enterprise Vault\FSA\PlaceholderService

Content:
DWORD

(Default value: 20)

Description:
You can specify a maximum rate of recall on each computer that runs a placeholder service, thus controlling the rate at which an individual user can recall files. By doing so, you also prevent any applications that do not honor the file system offline attribute from recalling all files that have been archived from a volume.

The default maximum rate is 20 recalls in 10 seconds. This recall limit applies to all users except, by default, members of the local administrator's group.

If the recall limit is exceeded, the application receives an Access Denied status. How this is displayed to the user depends on the individual application.

RecallLimitMaxRecalls specifies the maximum number of items that a user is allowed to recall in RecallLimitTimeInterval seconds. Note: This feature is not available when archiving from Network Appliance NAS devices.

Allowing administrators to bypass recall limits

Key name:
BypassRecallLimitsForAdmins

Location:
HKEY_LOCAL_MACHINE\SOFTWARE\KVS\Enterprise Vault\FSA\PlaceholderService

Content:
DWORD

Possible values:
0—(Default) Recall limits apply to administrators.
1—There are no recall limits for administrators.

Description:
You can specify a maximum rate of recall on each computer that runs a placeholder service, thus controlling the rate at which an individual user can recall files. By doing so, you also prevent any applications that do not honor the file system offline attribute from recalling all files that have been archived from a volume. This recall limit applies to all users except, by default, members of the local administrator's group.

BypassRecallLimitsForAdmins controls whether the recall limits apply to members of the local administrator's group on the server that is running the placeholder service.

Appendix B: Command-line utilities

FSARunNow utility

Use the FSARunNow utility to do the following:

- Start archiving a specified file server.
- Synchronize permissions for a specified file server.
- Prune the earlier versions of archived files until the required number of versions remains.

Note that you can create batch files containing the required FSARunNow commands and use Windows Task Scheduler to run the files when required.

To run FSARunNow:

1. Log on to any Enterprise Vault server as an account with local administrator permissions, such as the Vault Service account.
2. Open a Command Prompt window.
3. Navigate to the Enterprise Vault program folder (normally C:\Program Files\Enterprise Vault).
4. Run FSARunNow using the syntax described below, in one of the following three forms:

To start archiving a specified file server:

FSARunNow Archive <FileServerEntryId> [VolumeEntryId][Report | Normal]

To synchronize permissions for a specified file server:

FSARunNow Synchronize <FileServerEntryId>

To prune the earlier versions of archived files:

FSARunNow Prune <FileServerEntryId> [Report | Normal]

Report	Generates a report outlining the changes that FSARunNow would make if you were to run it in normal mode, without making those changes.
Normal	Specifies a normal mode run.

Examples:

To perform an archive run in report mode, type the following:
FSARunNow Archive 1D6D9206BFDBFB846B2E0F8135A1989331d100002example.server.local report

To perform a synchronizing run in normal mode, type the following:
FSARunNow Synchronize
1D6D9206BFDBFB846B2E0F8135A1989331d100002example.server.local normal

To perform a pruning run in report mode, type the following:
FSARunNow prune 1AD6297BC643DCC40A924CAB74D0BCDCE141000server.example.net report

FSAUtility

FSAUtility is a command-line utility with which you can do the following:

- Re-create the placeholders for archived files in their original location. (This may prove useful if you need to restore a file server to its original state or to synchronize the file server with the Enterprise Vault archive. If multiple versions of the same file exist in the archive, the utility creates a placeholder for the latest version only.)
- Re-create archive points on the original path.
- Restore some or all of the archived files to their original location or a new location.
- Move placeholders from one location to another location. The corresponding files in the archive are also moved to the destination folder. The destination archive can be in a different Enterprise Vault store.

Note: To optimize performance and prevent inconsistent behavior, stop the FSA archiving tasks on the target server before you run this utility.

Syntax:

FSAUtility <command> <command-specific parameters> [general parameters]

The parameters are as follows:

Command

-c	Re-create the placeholders for archived files.
-a	Re-create archive points.
-t	Restore some or all of the archived files.
-m	Move placeholders to a different location.

Parameters for -c (re-create placeholders for archived files)

-s <UNC path>	Specify the UNC path to the target folder or volume.
-D <mm-dd-yyyy>	Optional. Only re-create placeholders for the items archived after this date.

Parameter for -a (re-create archive points)

-s <UNC path>	Specify the UNC path to the target volume.

Parameters for -t (restore original files)

-s <UNC path>	Specify the UNC path to the source folder, volume, or file server.
-e <ext list>	Optional. Specify the file types to restore as a comma-separated list of file name extensions (for example: *.xls,*.doc,*.txt -). By default, the utility restores all file types (*.*).
-d <UNC path>	Optional. Specify the UNC path to the destination folder, volume, or file server for the restored files.
-D <mm-dd-yyyy>	Optional. Only restore files archived after this date.

Parameters for -m (move placeholders to a different location)

-s <UNC path>	Specify the UNC path to the source folder.
-d <UNC path>	Specify the UNC path to the destination folder.

Examples:

The following command, which is running in report mode, re-creates the placeholders for the folder \\myserver\users and generates a log file that lists both successful and failed operations:

FSAUtility -c -s \\myserver\users -l 0 -r

The following command restores the Word and Excel files in the folder \\myserver\users and generates a log file that lists both successful operations and failed operations:
FSAUtility -t -s \\myserver\users -e *.doc,*.xls -l 0

The following command moves the placeholders from the first folder to the second folder:
FSAUtility -m -s \\myserver\users -d \\sample\share

In this case, the log file lists failed operations only.

The following command restores the Word and Excel files for an entire file server:
FSAUtility -t -s \\myserver -e *.doc,*.xls -l 0

The following command re-creates the archive points for the folder \\myserver\users in report mode. In addition, the command reports all the archive points in subfolders of this folder:
FSAUtility –a –s \\myserver\users -r

ArchivePoints utility

The ArchivePoints utility provides a convenient means to create and manage archive points. These points mark the top of each folder structure that File System Archiving is to store in a single archive.

Syntax:

ArchivePoints <action> <archive point path share name> [subfolders|nosubfolders] [<Template XML File>]

where:

<action> specifies the action to perform. This can be one of the following:

Create	Creates archive points.
Delete	Deletes archive points.
Find	Lists all the archive points beneath the specified network share.
Read	Displays the contents of the archive points.
Update	Updates the archive points on the specified folders with the contents of <Template XML File>. <Template XML File> is a mandatory parameter when you use Update.

<archive point path share name> specifies the full UNC path to the network share to which the command applies.

subfolders|nosubfolders specifies whether the action applies to each subfolder in the specified network share. The default is nosubfolders.

<Template XML File> specifies an XML template file with which you can override the following attributes when creating an archive point:

name	The archive name.
description	The archive description.
owner	The archive billing owner.
indexDisabled	Whether to disable (True) or enable (False) indexing for the files in the network share. By default, indexing is enabled.
indexingLevel	The indexing level.
deleteExpiredItems	Whether to delete automatically expired items from the archive.
prefix	A prefix to add to the start of the archive name. The remainder of the archive name is taken either from the Name attribute, if present, or the folder name.

For example, the following overrides all the defaults when creating an archive point:

```
<archivePoint>
    <name>Newton archive</name>
    <description>Isaac Newton's User Archive</description>
    <owner>astronomy\newtoni</owner>
    <indexDisabled>False</indexDisabled>
    <indexingLevel>brief</indexingLevel>
    <deleteExpiredItems>false</deleteExpiredItems>
    <prefix>User </prefix>
</archivePoint>
```

Examples:

To create an archive point on folder \\myserver\users\jones, type the following:

ArchivePoints create \\myserver\users\jones

To list all archive points on share \\myserver\users, type the following:

ArchivePoints find \\myserver\users\jones

Appendix C: More information

Administrator documentation

The Enterprise Vault installation kit ships with comprehensive documentation.

Symantec Enterprise Vault—Introduction and Planning

Before deploying Enterprise Vault, it is necessary to understand the concepts and implications of archiving. This guide contains in-depth information about planning a successful deployment of Enterprise Vault.

Installing and Configuring Symantec Enterprise Vault

This document details the system preparation and installation process and helps in understanding the necessary steps for deploying Enterprise Vault.

Symantec Enterprise Vault Administrators Guide

This document contains detailed instructions about common tasks such as setting up roles for managing admin security or performing daily operations such as backing up Enterprise Vault.

Enterprise Vault release notes

The Enterprise Vault installation kit ships with a ReadMeFirst.htm document, which includes important tips, a description of the newest features, known issues, guidelines for improving performance, and an index of official product documentation.

Customer training course

Symantec Education Services offers a comprehensive Enterprise Vault training course. For more information on Symantec Education Services course offerings, please visit www.symantec.com/enterprise/training.

About Symantec

Symantec is a global leader in infrastructure software, enabling businesses and consumers to have confidence in a connected world. The company helps customers protect their infrastructure, information, and interactions by delivering software and services that address risks to security, availability, compliance, and performance. Headquartered in Cupertino, Calif., Symantec has operations in 40 countries. More information is available at www.symantec.com.

For specific country offices and contact numbers, please visit our Web site. For product information in the U.S., call toll-free 1 (800) 745 6054.

Symantec Corporation
World Headquarters
20330 Stevens Creek Boulevard
Cupertino, CA 95014 USA
+1 (408) 517 8000
1 (800) 721 3934
www.symantec.com

02/07 11859826