



Technical report: Symantec Enterprise Vault and IBM System Storage N series

Integration for file-system archival

• • • • • • • • •

Document NS3383-0

October 10, 2007



Table of contents

Abstract	3
Executive summary	3
Background	3
Introduction	4
Symantec e-mail archival.....	4
Symantec file-system archival	5
IBM N series Storage Solution	7
Network Connectivity.....	7
Microsoft System Environment.....	8
Prerequisites.....	8
Design Configuration	9
Configuration of Local Disks and SnapDrive.....	10
Mapping the Network Share on NearStore	11
SnapLock: Next Step toward Simplifying Compliance	11
Current configurations	13
Deployment overview.....	14
Enterprise Vault: Preinstallation	15
Installing the prerequisites	16
Configuring SQL Server Login	16
Configuring the Microsoft Message Queue Server	16
Completing the Preinstallation Tasks	16
Installing Enterprise Vault	18
File-system archival	19
Design and Sizing Requirements	19
Topology Selection.....	20
Vault Server Specification	20
Database Storage Configuration	20
Installing File-System Archival Component	21
Enterprise Vault Configuration	23
Adding Vault Services	24
Adding a File Server.....	28
Creating a New File Server Archiving Policy.....	29
Policy Creation for Volumes and Folders.....	31
Scheduling File System Archiving.....	31
Summary	34
Caveats	34
Trademarks and special notices	35



Abstract

In addition to e-mail archival, file-system archival offers a new feature to archive needed files and save storage space. Symantec Enterprise Vault has a compelling product to meet these challenges and improve the customer experience. This paper describes the method to integrate Symantec Enterprise Vault with an IBM System Storage N series configuration to take advantage of file-system archival and backup, recovery, and compliance solutions from IBM.

Executive summary

Messaging and collaboration servicing combined with an efficient archival solution in a configuration where a storage solution can improve the total customer experience is the secret to addressing customer problems in that space. Symantec has recently added the file-system archival feature in addition to e-mail archival in its Enterprise Vault product. This paper discusses the procedure to integrate Symantec products and their required components and other products with an IBM® System Storage™ N series solution. This paper will discuss the file-system archival (FSA) feature in an IBM N series environment.

Background

In the age of information, e-mail has become the mode of communication in the business community. Attachments to e-mail have served as a business solution in the communication process, and they are now a norm in the day-to-day business process. The increased acceptance has led to documents and messages being stored electronically. Once information is stored electronically, need for an efficient method of storing, archiving, and managing it is a challenge. In addition to e-mail archival, FSA offers a new feature to archive needed files and save storage space. Symantec has a compelling product to meet these challenges and increase the customer experience. In addition to e-mail archival, end users are looking for ways to manage file systems and ways to intelligently archive and manage files.

Several regulations have been enacted mandating compliance with content archiving specifications, while retaining the ability to produce the information when needed. Although not all businesses are required to follow these regulations, businesses are looking for ways to protect content proactively and cost-effectively, to reduce exposure, and streamline productivity. Regulations such as SEC Rule 17a-4, the Healthcare Insurance Portability and Accountability Act (HIPAA), and CFR 21 are forcing certain businesses, such as those in the financial, insurance, and healthcare sectors, to protect content. Other legislative initiatives are aimed at protecting personal data, such as social security numbers, and affect a wide range of companies.

In adopting policy to protect data, archiving and managing file systems have become important requirements. To solve the customer's business needs, Symantec has introduced a feature called file-system archival (FSA), beginning with its Enterprise Vault 5.0 Service Pack 3 product. Businesses require an efficient backup and restore method for file-system archival. IBM N series adds value to FSA in providing performance-improving backup, data replication, and data recovery features that come with its own operating system, called IBM System Storage N series with Data ONTAP®.

Introduction

Businesses need a plan to store and archive content, including files, that enables them to search quickly, yet also provides data security. Content has to be stored and managed, and users must be able to search the content and retrieve it as needed. Symantec Enterprise Vault has these capabilities. It is also important to understand the fundamentals of unstructured information lifecycle management to know business needs. Unstructured information includes a variety of file types and is spread throughout an organization (compared to structured information, such as database files, which typically are centralized and easier to protect and/or manage).

Enterprise Vault provides functionality and features to achieve complete business success. This paper will explain the basic procedure to integrate Enterprise Vault software with IBM N series storage solutions. It also discusses the FSA feature of Enterprise Vault in IBM N series configurations. The FSA feature allows archiving and managing the files on the file system, including private and public folders.

This paper will attempt to explain the steps involved with integrating Enterprise Vault with IBM N series storage. This paper does not discuss the benefits of file archival, performance, backup, recovery, and configurations to include disaster recovery by deploying the data replication features of IBM N series features.

Symantec e-mail archival

Developing software to archive content by optimizing the use of storage, yet providing management simplification was a challenge. Symantec has developed a suite of products to effectively address the above-mentioned issue and provide a method for quick implementation, a system that is transparent to users with an open application programmable interface (API). Symantec offers an Enterprise Vault product that provides store, index, search, and retrieve capabilities in Microsoft Windows® Exchange environments. The FSA feature in Enterprise Vault enables file archival and search and retrieval capabilities. Symantec provides open APIs to store, manage, and discover any content in a customer's environment.

Symantec Enterprise Vault works in Windows Exchange environments, and it is helpful in understanding the need for Enterprise Vault where mailboxes store the data. Enterprise Vault allows storing content as standard files, one file per message in one or more NT File System (NTFS) partitions. This approach will offer extensive benefits such as the following:

- No additional maintenance is required
- It is easy to recover a corrupted item as compared to the entire database
- More data can be stored without performance degradation
- Central management of storage for efficient usage of storage
- Support for multiple Exchange Servers.

Figure 1, on the following page, shows a simple configuration of Symantec Enterprise Vault managing e-mail and file-system archival.

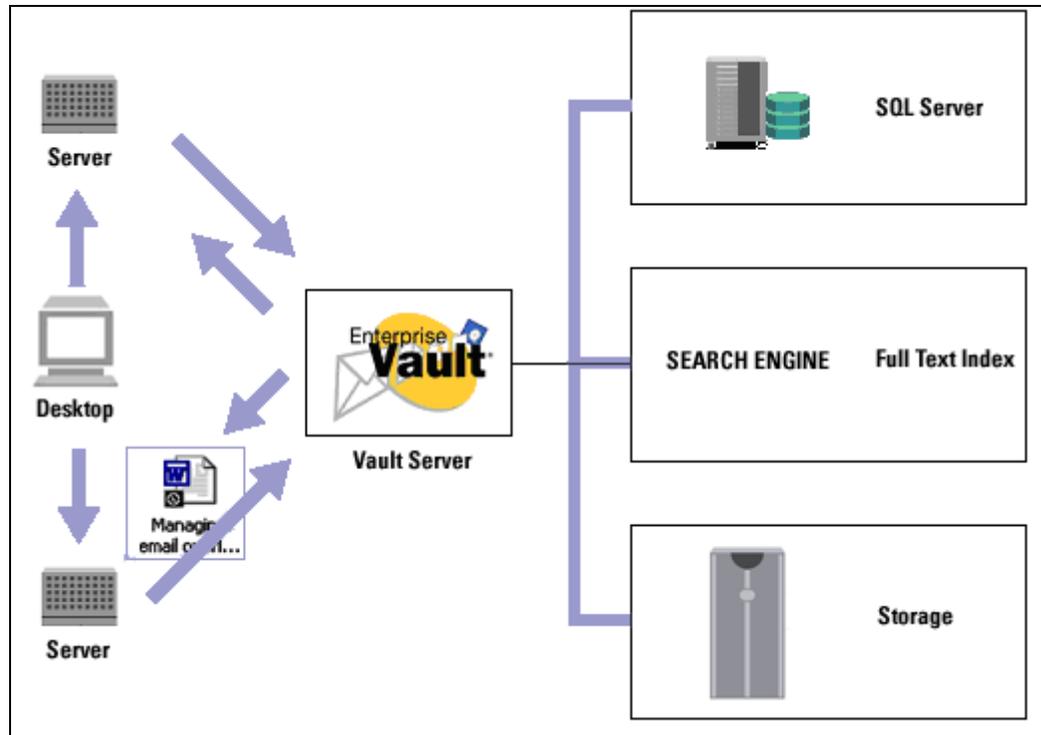


Figure 1. Symantec Enterprise Vault environment configuration.

Symantec file-system archival

Symantec has extended Enterprise Vault functionality beyond e-mail to file-based data. It includes the capability to archive the system files. It is also possible to configure Enterprise Vault to use the FSA component. FSA functionality is limited to file systems that can be presented as NTFS. Files that reside on network share, corporate data, and documents are typical examples for archiving into Enterprise Vault. Note that it is dangerous to archive Windows system files, and this paper strongly recommends excluding the archival of system files and other files that are critical to running the operating system. This paper recommends avoiding archiving system-related files of the operating system. In such scenarios, the system has to be recovered using the system backup, and Enterprise Vault data has to be recovered from its backup.

The creator and migrator tools in Enterprise Vault allow moving archived data to a different device. By using FSA, files can be stored where it is easier to manage the data to take care of backup/recovery and data replication. Archived files provide a centralized location for information storage and allow data mining. FSA provides flexible policy control and quick recovery of data.

It is important to note that Enterprise Vault allows backup of content that is required to be stored with lifecycle management and provides end-user access to the data. The Enterprise Vault file system adds value by addressing data management and maintains end-user access. Flexible policies can be set for specified types of files to be archived and managed. As depicted in Figure 2, a single archive for all data is allowed without the restriction of files, electronic content, or other content type.

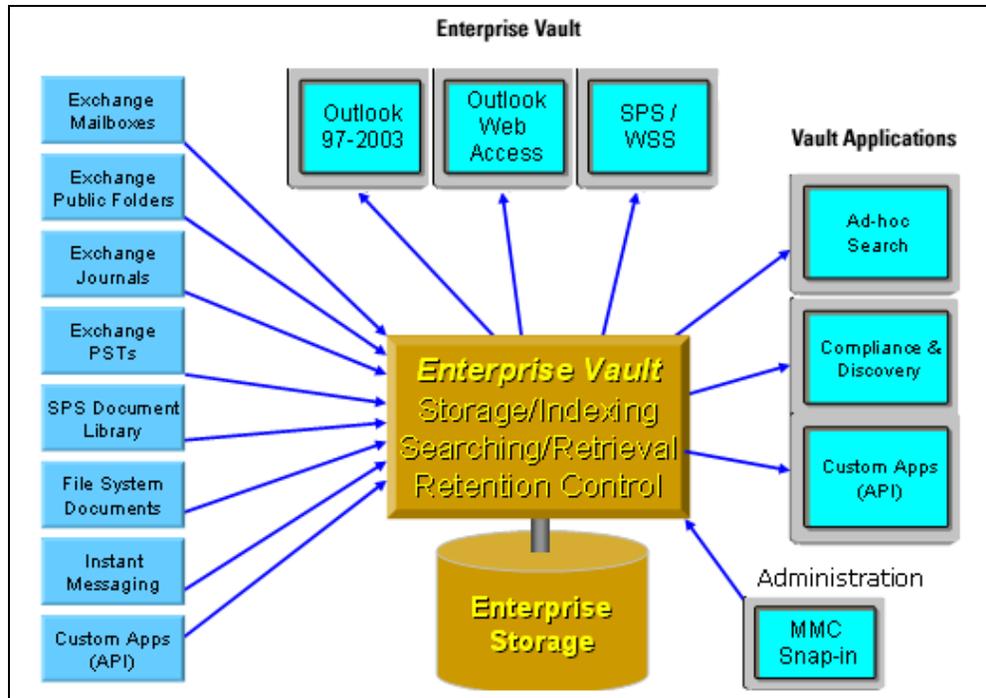


Figure 2. Symantec Enterprise Vault overview.

Figure 3 shows a framework for Enterprise Vault highlighting all the components present in an archiving environment. Note that the file system component is one level below the universal access layer, along with Exchange and SharePoint components. This shows that FSA can be configured with or without other components such as Exchange Server or SharePoint product features.

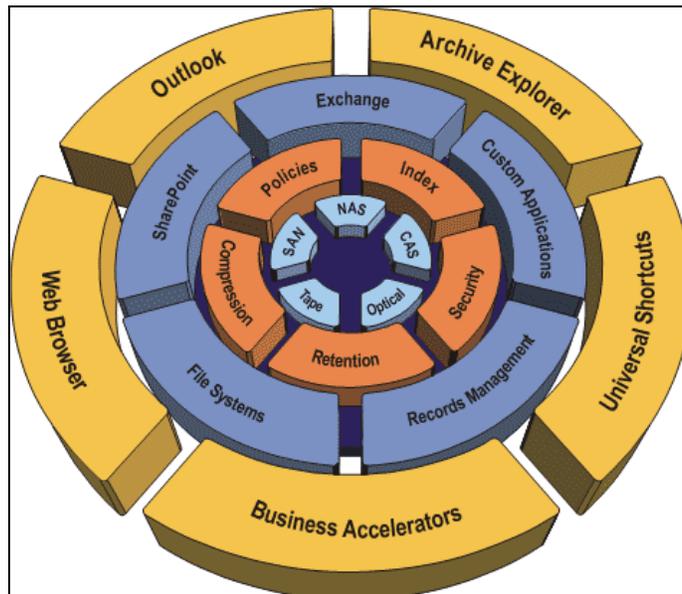


Figure 3. Enterprise Vault: showing the file-system archival component.



IBM N series Storage Solution

IBM N series storage devices contain many redundant hardware features. Built-in RAID (redundant array of independent disks) protects against downtime due to disk failures. In the event of a disk failure, automatic reconstruction takes place on a hot spare disk with notification sent to the system administrator. The Data ONTAP software storage health monitor proactively monitors the disk drives and storage connections for any potential problems. Redundant power supplies and cooling fans are included in the system unit and disk shelves. Cooling fan speed and system temperature are also monitored, and notification is sent if there is a problem. Disk drives, power supplies, and cooling fans are all hot swappable. Additional hardware redundancy can be achieved by deploying the following configurations: virtual network interfaces, Fibre Channel (FC) multipath, and IBM N series storage cluster configurations.

It has become standard for enterprise-level customers to configure a private Gigabit network connection between the Enterprise Vault server, Microsoft® SQL Server, and IBM N series storage devices. However, in FSA, local clients on the corporate network infrastructure may also connect to exploit the benefits of FSA.

IBM N series systems, including IBM System Storage N series with NearStore® feature, support the configuration of different protocols depending upon application requirements. IBM System Storage N series with SnapLock® offers the benefits of non-erasable, non-rewritable storage of write once, read many (WORM) volumes. Configuring and using SnapLock requires enabling common internet file system (CIFS) services on the IBM N series devices. The CIFS protocol uses general network connections to transfer data. To install and use SQL Server, this paper recommends using Internet Protocol Small Computer System Interface (iSCSI) or Fibre Channel Protocol (FCP). FCP and iSCSI allow local disks to be set up on the Windows Server. IBM System Storage N series with SnapDrive® software will ease the administration and management of local configured disks. SnapDrive also helps to scale disk size as data grows and helps to manage the backup and recovery features.

IBM N series with Data ONTAP 7G allows storage administrators to create more flexible and scalable storage configurations by using features such as IBM System Storage N series with FlexVol®, where the volume can be grown or shrunk as needed on IBM N series storage.

Network Connectivity

To install and configure Symantec Enterprise Vault, the Microsoft Exchange Server SQL Server database requires different types of network connections. Network connection via FCP or iSCSI between Exchange Server and IBM N series storage is recommended. The storage configured with either iSCSI or FCP can be maintained much more easily with a software feature called SnapDrive from SnapDrive will ease storage device maintenance with data management for backup/recovery or scaling the storage devices as data grows. Providing continuous access to storage is one element of high data availability. The other elements are the integrity and recoverability of the data. The IBM N series storage appliance has several built-in features and optional software for data integrity and protection.

Enterprise Vault also requires a network connection between the Windows Servers and the IBM N series storage devices. The network connection between the Windows servers and the IBM N series device must use Gigabit Ethernet.



It has become standard to use a private Gigabit Ethernet network connection between the Symantec Enterprise Vault server, the Exchange Server, the Active Directory domain controller server, and the IBM N series data devices. Either setting up a separate switch or creating a virtual LAN (VLAN) on existing switches would suffice equally. Client connectivity to Exchange Server can continue over the current network infrastructure.

FSA allows setting a policy for private and public folders and hence using the network shares. Having this archival available on a shared network drastically increases data storage usage. FSA helps to archive the files and to index the content, and hence the data mining becomes an important benefit.

Microsoft System Environment

Symantec Enterprise Vault works on the Windows platform and supports Microsoft Exchange Server and SQL Server relational databases on Windows 2000, Windows XP, and Windows 2003 Server platforms. Microsoft Exchange requires the local disk configuration, and hence the use of iSCSI and/or FCP local disk configurations is required on IBM N series devices. To ease storage management issues, SnapDrive software to be installed and configured on both the Exchange and SQL database server. Enterprise Vault supports IBM N series storage as long as it can be presented to the system as an NTFS file system. Configuring the NTFS system using IBM N series storage can be achieved using the CIFS, iSCSI, or FCP protocols.

Prerequisites

Before proceeding with the installation of Enterprise Vault software, complete the prerequisite worksheet. This will help you install and configure the Enterprise Vault Server successfully. Refer to the Enterprise Vault product manual to verify the prerequisites for installing the Enterprise Vault server.

Enterprise Vault supports only a SQL Server environment, and SQL Server is required on a Windows Server. In a production environment, Exchange Server and SQL Server installed on separate Windows Servers. This will provide much-required performance when dealing with large amounts of data.

Enterprise Vault requires network share to the storage front to enable searching and archiving files. The network share should use the Unified Network Connectivity (UNC) path to maintain the same network path across multiple client machines and servers. If the storage is mapped using network share and assigned a drive letter, it may create issues for data visibility across different client machines, and the drive letter may differ and cause some issues.

Regarding the Windows Server requirement, a separate server is available for each Exchange Server and database server and all the required Windows service packs installed. If the FSA component configured as a standalone feature, Exchange Server is not required. However, most customers use an Exchange Server configuration, and hence this paper assumes that Exchange Server is used.

The IBM N series storage requirement depends on the configuration of SQL Server and the Enterprise Vault software and the storage requirement of vault stores. Typically, Exchange Server, Enterprise Vault software, and SQL Server are configured on high-performance storage, while



NearStore is used for deployment for storing data on vault stores. Filer configuration requires iSCSI and/or FCP to complete the local storage requirements.

Note that IBM N series with Data ONTAP 7.1 and later releases of this operating system, support Enterprise Vault as long as the necessary protocol is enabled. With these releases, it is possible to delete the retention-expired files in the vault server. IBM N series supports storage of non-erasable, non-rewritable WORM volumes enabled by SnapLock starting from Data ONTAP 7.1.

The hardware requirement for Enterprise Vault may vary largely depending on the configuration, such as the number of Exchange Servers, archival destination, amount of data in terms of bandwidth, and storage media. However, if only FSA component is configured without installing Exchange Server, the configuration will change significantly, as we have to focus on the number of files, file size, and mode of archival, etc.

Enterprise Vault must install and configure the Windows Server with Internet Information Services (IIS) and Microsoft Message Queue Server (MSMQ), and the services started. The software configuration also requires Microsoft SQL Server. Even though the Microsoft Exchange Server is not required to install and configure FSA, Outlook 2000 or Outlook 2003 installation is required, as Enterprise Vault uses messaging application programming interface (MAPI) services. Symantec software such as Enterprise Vault redistributable software media comes with the required Windows operating system prerequisite software.

Design Configuration

Sections 3.1 and 3.2 gave a brief product overview of Enterprise Vault. A Symantec Enterprise Vault system grouped into four sections: information source agents, user clients, and vault applications together with Enterprise Vault to store and administer the vault. Information agents include Exchange mailboxes, public folders, file system, SharePoint, etc., and vault administration will be done with Microsoft Management Console (MMC).

In a Symantec environment, several message servers and desktop clients may exist. Enterprise Vault supports multiple Exchange Servers, and a typical limit is eight Exchange Servers per Enterprise Vault. Each Exchange Server configured is connected with a single SQL Server. IBM N series systems are configured with storage area network (SAN) or internet protocol (IP)-based SAN configuration. Primary storage is used to configure the Exchange Server and SQL Server data. Note that the same system can be configured on both Exchange Server and the database server as local disks to install and configure the Exchange Server and SQL Server database.

To archive e-mail and files using FSA, a second IBM N series with NearStore feature is configured using the network path. For this purpose, we used the CIFS protocol to map the drive using the UNC path. Providing an UNC path provides the same path name across the network.

In Figure 4, Enterprise Vault has access to both Exchange Server and SQL Server database data. The system is configured to have a local storage configuration, and NearStore provides a network share to be able to archive the files and e-mail.

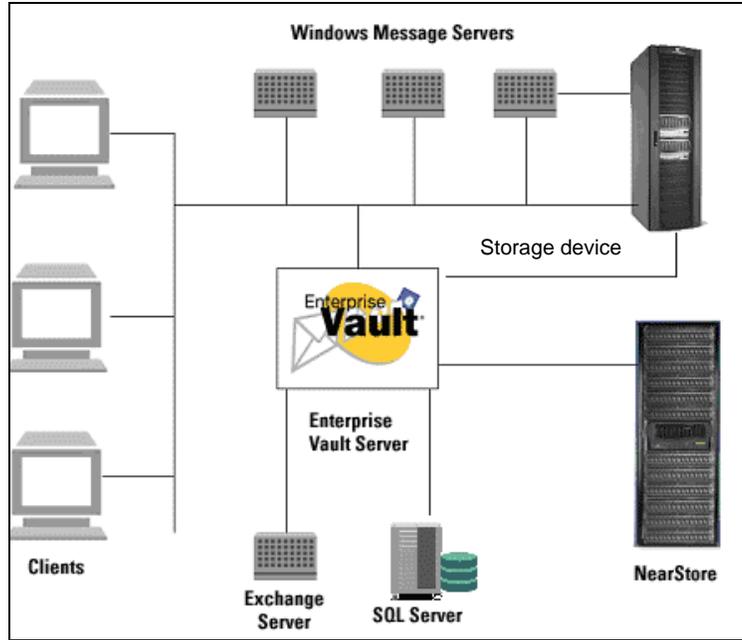


Figure 4. Enterprise Vault design configuration.

Configuration of Local Disks and SnapDrive

SnapDrive software integrates with the Windows Volume Manager so that the IBM N series can serve as virtual storage devices for application data in Windows 2000 Server and Windows Server 2003 environments.

SnapDrive manages virtual disk logical unit numbers (LUNs) in an IBM N series, making these virtual disks available as local disks on Windows hosts. This allows Windows hosts to interact with the virtual disks just as if they belonged to a directly attached RAID array.

SnapDrive enables online storage configuration, virtual disk expansion, and streamlined management. It integrates IBM System Storage N series with Snapshot™ technology, which creates point-in-time images of data stored on virtual disks. It works in conjunction with IBM System Storage N series with SnapMirror® software to facilitate disaster recovery from asynchronously mirrored destination volumes.

SnapDrive supports both iSCSI and FCP, and using either an iSCSI software initiator or host attach kits, SAN or IP-based SAN configurations can be configured on IBM N series storage. To install SnapDrive software on IBM N series storage devices, refer to the appropriate SnapDrive installation and configuration guide.

The recommendation is that any system connected to a host should reside in the same broadcast domain as that host, so that virtual disk I/O commands do not need to traverse router hops. For Windows cluster configurations, do not permit internal cluster traffic on a Gigabit Ethernet (GbE) network used for host-system data transfer. Instead, use a Fast Ethernet connection for all cluster traffic. This practice ensures that a single network error cannot affect both the connection for internal cluster traffic and the connection to the quorum disk.



Mapping the Network Share on NearStore

In order to provide the archival destination, network connectivity between the Enterprise Vault server and IBM N series with NearStore feature must be configured. To complete the configuration, verify that the NearStore server name is entered in the Windows domain and the network connectivity is established. Once the network connectivity is established, create the volume of desired size. Starting with Data ONTAP 7.1, IBM N series provides great flexibility in defining and configuring volume sizes. Depending on the need and growth of data, the volumes can either be expanded or shrunk provided the right type of volumes is created, such as FlexVol, which allows growing or shrinking the disk volume size.

SnapLock: Next Step toward Simplifying Compliance

SnapLock technology has been added as a part of the Data ONTAP operating system, and it runs on currently supported IBM N series storage systems. This includes both IBM N series primary and nearline storage systems (IBM N series with NearStore feature). SnapLock is an integral part of the operating system. Hence, no separate installation is required to install SnapLock. By entering the necessary license key, the SnapLock feature provides the business with a simple, yet robust technology. SnapLock software feature allows to quickly deploying non-erasable, non-rewritable magnetic storage media. By providing such a feature, it allows the customer to exploit the benefits offered by IBM N series storage systems such as instant backup/quick recovery and an efficient and fast way to replicate data to a different location.

Regulations to protect content for a certain period require regulated industries such as financial services, healthcare, and utilities. SnapLock enables WORM and non-WORM storage volumes to coexist at the same time. This means that SnapLock volumes can coexist with volumes other than SnapLock within one system. Each SnapLock volume will have a unique volume identifier to distinguish it. A unique volume identifier (UUID) is associated with a SnapLock volume. This will help to maintain the unique identification of the SnapLock volume for future retrieval requirements. SnapLock is available on all IBM N series systems, and it utilizes standard CIFS and NFS protocols to store and access files. Volumes create command, "vol create -L," creates a SnapLock volume. Verify that the new volume created is in fact a SnapLock volume by using the "vol status" command. There are two distinct types of SnapLock volumes: compliance and enterprise. This depends on the type of SnapLock license on the IBM N series storage device. Two types of SnapLock volume are available, the regular SnapLock Compliance and SnapLock Enterprise. The IBM N series storage system prevents removal of a SnapLock Compliance volume before the expiration of retention files on the volume. Certain users, for example, system administrators, may destroy SnapLock enterprise volumes before the expiration of their retention period. Certain prerequisite software components such as SnapDrive are required on IBM N series devices. An example of such a requirement would include CIFS, iSCSI, or FCP and a license for installing SnapDrive software.

The process of committing files (data) is similar to the process of writing to optical platters. Once the file (data) is committed to non-erasable, non-rewritable status, SnapLock will set the WORM volume attributes with the set retention properties for archival purposes. The implementation procedure is the same for both SnapLock compliance and the enterprise volumes. In order to meet the different requirements, administrators are allowed to create single or multiple SnapLock volumes on the same IBM N series storage system.



SnapLock compliance configuration is targeted towards business applications that are required by law/regulations to archive on non-erasable, non-rewritable storage media. The files committed to a SnapLock volume will retain the file attributes until the expiration of the retention period. This means the volume will remain at least until all files reach their retention expiration period.

Currently IBM N series storage devices support creating a flexible volume depending on the storage requirement by configuring a larger aggregate to allow the creation of smaller and flexible volumes to be created on top of the aggregate. This will enable storage administrators to exploit the benefits of storage provisioning.

IBM N series offers a more robust solution with a new type of RAID protection named IBM System Storage N series with RAID-DP™. RAID-DP stands for RAID Double Parity, and it significantly increases the fault tolerance from failed disk drives over traditional RAID. At the most basic layer, RAID-DP adds a second parity disk to each RAID group in a volume. Whereas the parity disk in a RAID4 volume stores row parity across the disks in a RAID4 group, the additional RAID-DP parity disk stores diagonal parity across the disks in a RAID-DP group. With these two parity stripes in RAID-DP, one horizontal and the other diagonal, data protection is obtained even in the event of two disk drives failing in the same RAID group.

On our system, the "vol status" command displayed the following command (where "EV" stands for Enterprise Vault).

```
boy> vol status
      Volume State      Status      Options
      vol0 online      raid4, trad root
      voll online      raid_dp, flex create_ucose=on,
                        convert_ucose=on
      EV online        raid_dp, flex create_ucose=on,
                        convert_ucose=on
      SLE online        raid_dp, trad no_atime_update=on,
                        raidsize=14,
                        snaplock_enterprise
```

If the SnapLock volume is not available, create a non-erasable, non-rewritable, WORM-enabled volume using an IBM N series storage device command-line interface by issuing a "vol create -L" command. On our system, a SnapLock volume, "FSALOCK," was created with a RAID-DP group. After creating "FSALOCK," the "vol status" command displayed the following output:

```
boy> vol status
      Volume State      Status      Options
      vol0 online      raid4, trad root
      voll online      raid_dp, flex create_ucose=on,
                        convert_ucose=on
      EV online        raid_dp, flex create_ucose=on,
                        convert_ucose=on
      SLE online        raid_dp, trad no_atime_update=on,
                        raidsize=14,
                        snaplock_enterprise
      FSALOCK online    raid_dp, trad no_atime_update=on,
                        snaplock_enterprise
```



After creating the volume, the next action is to create the needed qtrees on SnapLock and on our system; we created a qtree called "BusinessApp." The following output displays a sample output for creating a qtree and a CIFS share:

```
boy> qtree create /vol/FSALOCK/BusinessApp mixed
boy>
boy> cifs shares -add BA /vol/FSALOCK/BusinessApp
boy>
```

The IBM System Storage N series with FilerView[®] utility also allows creating qtrees and CIFS shares. However, the command-line interface is used to create a SnapLock volume.

It is also an important requirement to enable and set the IBM System Storage N series with ComplianceClock[™]. Unless ComplianceClock is initialized, expired files cannot be deleted. ComplianceClock can be initialized only once for a particular IBM N series storage device.

ComplianceClock is initialized only once for each unit in its lifetime. This paper suggests verifying the time set and continuing with ComplianceClock initialization as follows:

```
Filer>date -c initialize
```

ComplianceClock is supported in IBM N series with Data ONTAP 7.1 and later releases.

Current configurations

In this paper, we used a simple configuration with Enterprise Vault on a Windows Server. SQL Server is running on a separate Windows Server. In our setup, the filer configuration included SnapDrive software to set up and manage the local storage. The Enterprise Vault server and SQL Server software products were installed on the configured local disks. An IBM N series with NearStore feature device may be configured as a destination for archival of files. No Exchange Server was installed in this configuration.

Deployment overview

This paper discusses FSA and skips the deployment of Exchange Server. FSA is included in the Enterprise Vault architecture with Microsoft Exchange Server. This paper will focus on deploying Enterprise Vault for enabling the FSA component with IBM N series storage devices. The setup is used in this paper as an example and a starting point. See Figure 5 for an overview.

IBM N series storage systems include the primary device and NearStore feature for nearline storage purposes. IBM N series offers higher performance storage systems that are highly suitable for running enterprise-level databases and other applications. The IBM N series with NearStore feature is a disk-based nearline storage system with inexpensive ATA attachment (ATA) disk drives running on the Data ONTAP operating system. In addition to cost-effective advantage, IBM N series with NearStore devices offer much required data protection, instant backup and recovery, and efficient data replication features. An IBM N series with NearStore feature device can scale from 7TB to 504TB.

Today's business needs require more and more data to be archived and accessed less frequently. Unlike tape backup, the IBM N series with NearStore feature offers quick recovery of data as and when required. Less frequently accessed files are archived onto the nearline device. Both filer and nearline devices support non-erasable, non-rewritable, WORM volume capabilities. This feature helps Enterprise Vault users to use the retention categories to set the archival destination and retention policy on IBM N series with SnapLock volumes.

In the absence of a higher performing system, SQL Server can be installed on an IBM N series with NearStore feature. If SQL Server is being installed on an IBM N series with NearStore feature, this paper recommends installing SnapDrive and iSCSI or FCP to configure the local disks. Running SQL Server using CIFS is known to cause certain issues, and hence it is not recommended.

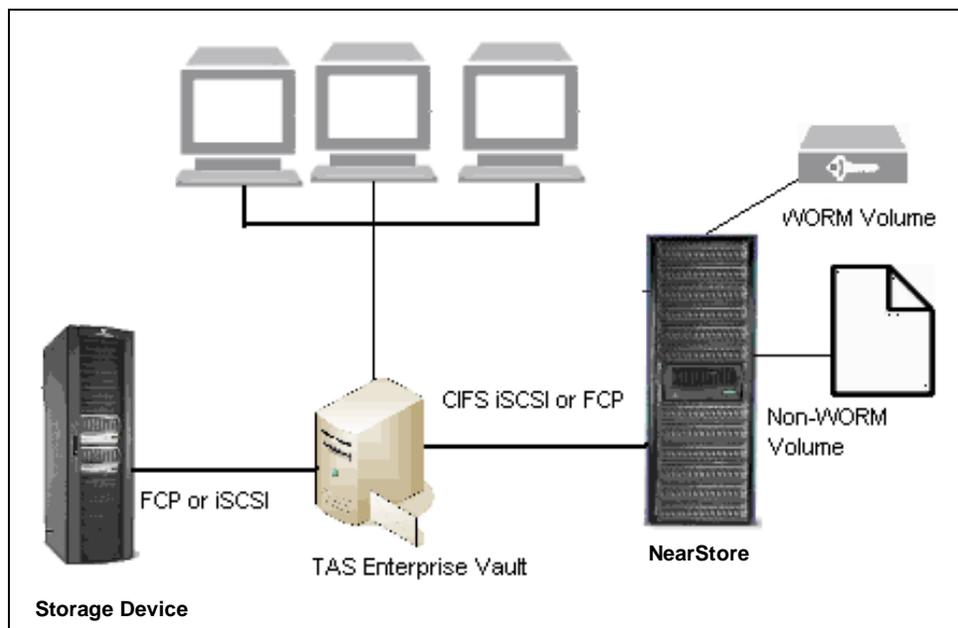


Figure 5. Enterprise Vault deployment overview.



Enterprise Vault: Preinstallation

It is important to understand the different components that are installed with Enterprise Vault. This section will provide the information required for preinstallation, including the software required and the tasks that are to be performed before installing Enterprise Vault.

Enterprise Vault has the following components.

- Windows services
- Vault administration console
- Web-based components to provide access to archives
- User extensions to allow clients to access archived items.

Preinstallation tasks include analyzing the requirement for performance and high availability, and the existing infrastructure architecture helps to implement a right solution. If FSA is installed without Microsoft Exchange Server, configuration details regarding Exchange Server may be skipped. Enterprise Vault has several services, and installing these services will enable the vault administrator to configure and run services on that server. The service components can be installed on any computer on which the services are run. Some of the services included in Enterprise Vault are the following:

- Admin service. One per server, and this service is installed automatically with the installation of any of the service components of Enterprise Vault.
- Indexing service. One per server, and it must have a connectivity to a physical storage location to store the index data.
- Storage service. One per server; requires connectivity to a physical storage configuration and access to Microsoft SQL Server for the vault store databases, IIS and MSMQ services.
- Shopping service. Requires a physical connectivity to storage configurations and IIS.
- Public folder service. Required for each public folder root directory.
- Journaling service. Required for each journal mailbox. It requires MSMQ and Collaboration Data Object (CDO) to run.
- Retrieval service. Required for each Exchange Server, and it should be installed on the same Enterprise Vault server as the archiving or journaling service.

Verify that Active Server Pages, IIS, Microsoft .NET, and MAC components are installed and registered as a virtual directory in IIS called "Enterprise Vault." Exchange forms are installed in the Exchange organizational forms library with ownership rights to that library. The vault administration console is a snap-in to the MMC. The administration console may be installed on any computer from which Enterprise Vault is to be managed.

Installing the prerequisites

In order to have a successful installation and configuration of Enterprise Vault, it is suggested that you follow the prerequisite software sequence to avoid issues with dynamic library loads (DLLs). Here is the sequence to complete the preinstallation tasks. It is important to note the platform configuration where the installation and configuration needs to be completed. If you are installing Enterprise Vault on Windows Server 2003 and Windows 2000 communicating with Exchange Server 2003 or Exchange 2000 Server, follow these tasks in the order they are given.

Obtain all the prerequisite software and note the requirements for each service to be installed. Some services require a physical connectivity and running services such as MSMQ and IIS.

1. Windows 2003 or Windows 2000 with Service Pack 3-Windows 2003 Standard Edition or Enterprise Edition, Windows 2000 Advanced Server, or Windows DataCenter Server may be used.
2. Outlook 2000, which needs CDO components if Exchange Server is not running on the computer.
3. Install SQL Server; it is recommended that SQL Server be installed on a separate server. Enterprise Vault works with Windows authentication mode and mixed-mode authentication, and SQL Server must be case insensitive.
4. On Windows 2000 Server, verify that Service Pack 3 is installed.
5. MSMXL, which comes with the redistributable software folder on the Enterprise Vault software media. Alternatively, install Internet Explorer V6.
6. Microsoft Data Access Component (MDAC) V2.6 or later and the software that comes with the Enterprise Vault media.
7. Microsoft .NET Framework V1.1 software, which comes with the redistributable software folder on the Enterprise Vault media.

Enterprise Vault services need access to the network with appropriate access permissions. This is done with one service account. Enterprise Vault services run under this account, and it is shared by all Enterprise Vault computers in all Enterprise Vault sites. This account must be able to log on as a service and act as part of the operating system and debug program user rights.

Configuring SQL Server Login

To create the directory and vault databases, Enterprise Vault needs to access SQL Server. This means that before installing Enterprise Vault, you should verify the network connection between the Enterprise Vault server and the SQL Server machines. Then create a SQL login.

Configuring the Microsoft Message Queue Server

MSMQ Server must be configured on the Enterprise Vault server. The use of a simple domain name system (DNS) alias such as vaultserver.mydomain.com is advised.

Completing the Preinstallation Tasks

As a part of the preinstallation tasks, create a DNS alias; on our configuration, the alias evault1 name was used.

Creating a SQL Login

Using the SQL Enterprise Manager, create a SQL login for the vault service account. If a separate group manages SQL Server, contact the SQL database system administrator to perform the task. To create the required SQL login, use the SQL Enterprise Manager. On our system, we added a new login and grant access with the server role as a database creator was configured. Figure 6 shows a sample output for the properties of a new SQL login.

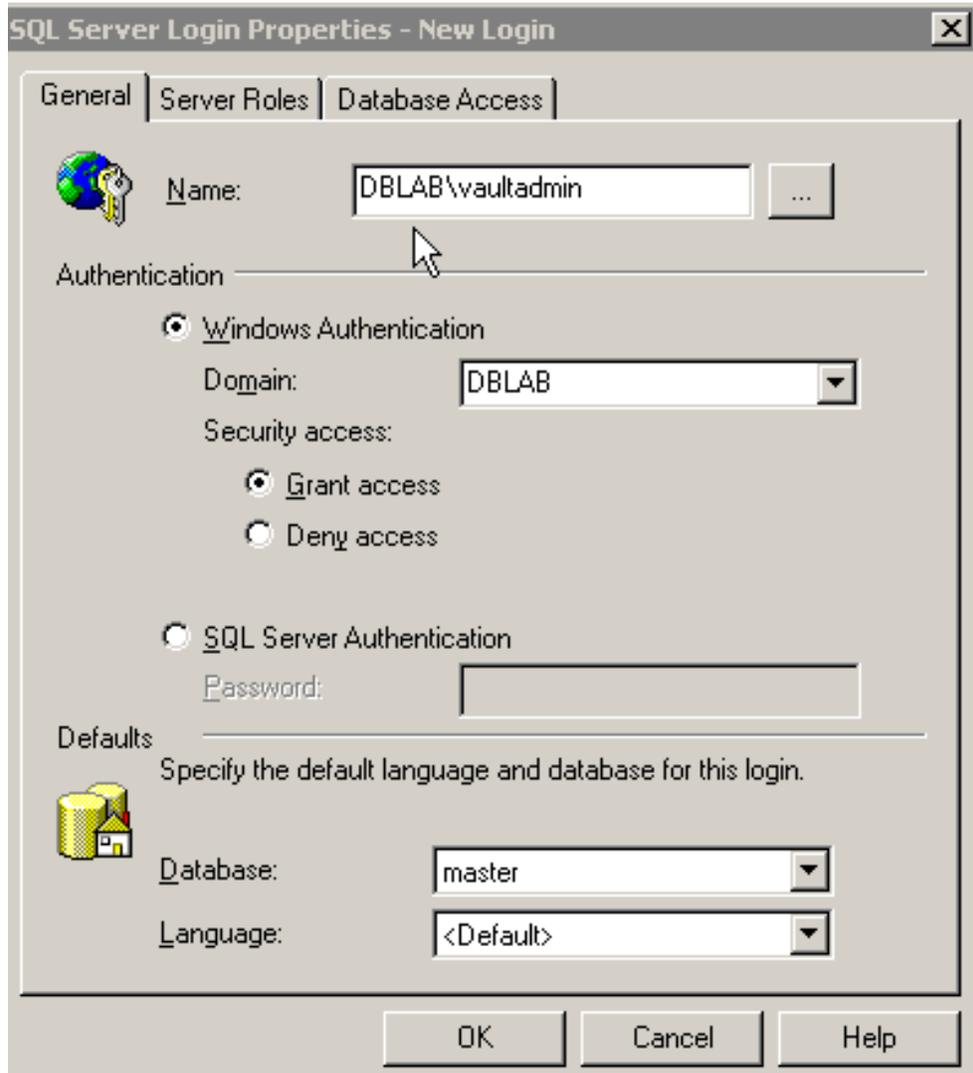


Figure 6. Creating a SQL login.

After creating the SQL login, verify that administrator privileges for the vault service account are set properly by opening MyComputer and manage local users and groups. Now the configuration is ready to install Enterprise Vault.

Installing Enterprise Vault

The previous section discussed preinstallation requirements. Once the preinstallation requirements are met, Enterprise Vault software can be installed. The Enterprise Vault installation process provides available choices for installing the required components (see Figure 7). Note that the administration console service is installed as part of Enterprise Vault installation. A virtual directory is created and registered in IIS called "Enterprise Vault." Before installing Enterprise Vault, stop the IIS admin to stop the dependent services and continue with the setup wizard, which will guide you through the installation by selecting Enterprise Vault and the administration console as the required components. Before installing Enterprise Vault, stop the IIS and follow the instructions to complete the installation. Note that FSA will be installed after the Enterprise Vault server is installed and configured, and hence the file placeholder services component may be unchecked.

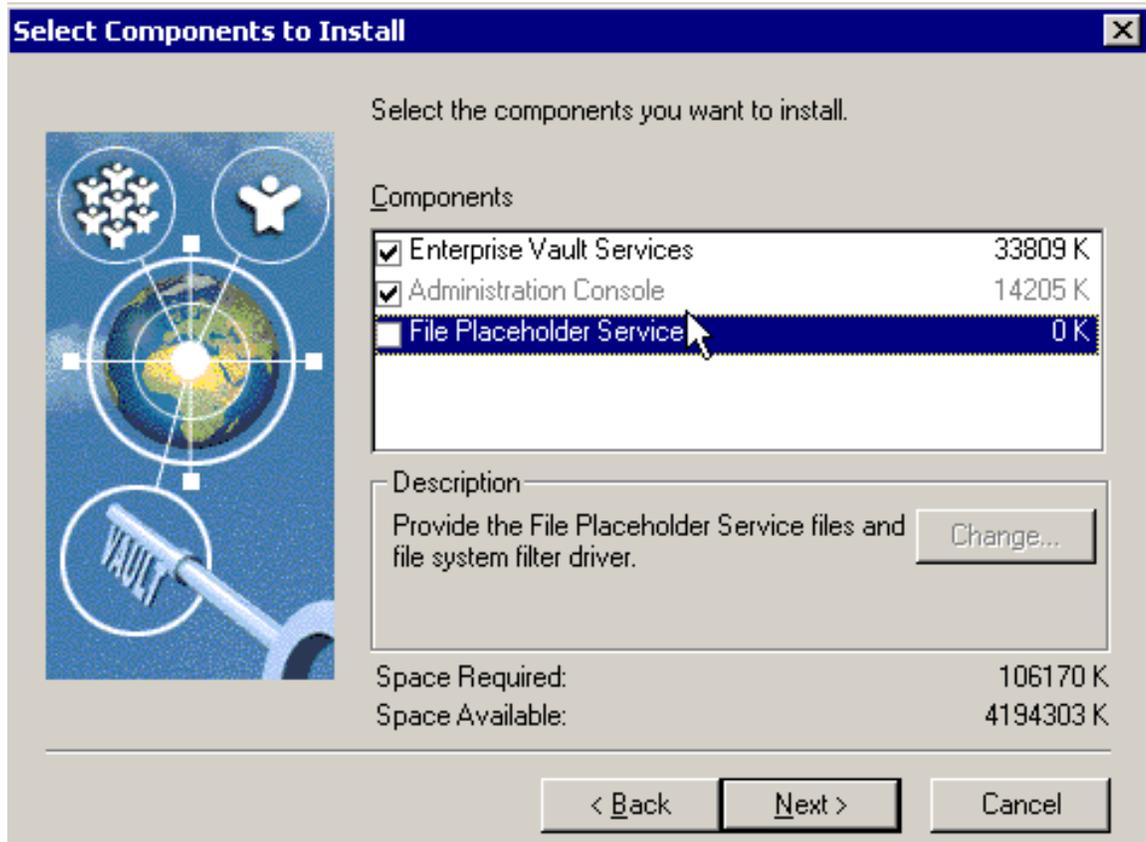


Figure 7. Enterprise Vault components.

When Exchange Server is not installed, the Enterprise Vault installation will display an error about the requirement for Exchange Server. Ignore the message and continue with the installation.



File-system archival

Enterprise Vault's information management functionality is extended to file-based data by providing FSA. FSA can be used along with different types of archiving such as e-mail or on its own. FSA is supported on all storage media, including IBM N series storage systems that have the capability to present it as NTFS devices. Files shared across the network and corporate data on a particular server are ideal candidates for archiving to a single location by using FSA. Using built-in Enterprise Vault tools such as collector and migrator, the archived data can be moved to a different location.

FSA offers several advantages by allowing the management of unstructured data from creation to destruction. It provides centralized repositories. More importantly, it provides indexing services, and hence content search is made easy. Unlike backup utilities, FSA allows complete control over data and lifecycle management. It maintains end-user access to archived data. Using Enterprise Vault server, flexible policies for archiving are set. Flexible FSA is a separately licensed component of the product, and hence no Exchange Server is required to install and configure Exchange Server.

Before installing FSA, analyze the method that can be related to the Exchange Server organization. Several possible installation strategies are available; a few common ones are described below

- One Enterprise Vault site for each Exchange Server site
- One Enterprise Vault site for a part of a single Exchange Server site
- One Enterprise Vault site for parts of multiple Exchange Server sites
- One Enterprise Vault site for many Exchange Server sites
- Several Enterprise Vault sites for one Microsoft Exchange Server site
- One Enterprise Vault site with no Exchange Server (this paper focuses on this model).

Design and Sizing Requirements

In order to design a proper FSA system, a reasonably good estimate of the server and storage requirements is required. The main design points should be based on a centralized, decentralized, or multiple-vault solution. It is also important to check what components of Enterprise Vault need to be archived. Some examples are given in the following list:

- Mailbox archiving
- Public folder archiving
- Journal archiving
- PST migration
- Office vault
- Compliance acceleration
- Discovery acceleration.

In the design analysis phase, consider if the high-availability feature is required. The high-availability solution with Enterprise Vault is designed considering the Enterprise Vault server availability rather than increasing the Enterprise Vault server performance. The high-availability solution may degrade performance while improving server availability in case of the failure of the primary Enterprise Vault server.

Lastly, consider how many vault servers need to be configured for an estimated of two years of storage. By following these requirements and analysis, a better system may be configured.



Topology Selection

In a decentralized design, vault servers are located at each site and use MSMQ services to connect remote vault servers over WAN. This leads to an advantage of MSMQ in reducing the traffic of MAPI requests. The topology adds complexity to the system design, and it may increase the access time to view the archived items.

The number of archiving services on a vault server depends on the limitations of Windows memory management that is available for Windows services under the local system account. With archiving and retrieval services configured, the same vault server can archive up to 14 Exchange Servers. This leads to the requirement of having multiple vault servers even when the archiving throughput could be achieved with a single vault server.

Vault Server Specification

Any Windows Server running Windows 2000, Windows 2003, or Windows XP is supported, and it is recommended to have dual Pentium or XEON processors with at least 2GB of memory to be able to handle archive requests. IBM N series storage provides the required RAID to protect data at the storage level. Configure sufficient storage for the SQL Server database data and logs. It pays well to design a simple configuration. Use of multiple servers and a topology other than centralized may not be a requirement in most scenarios.

Database Storage Configuration

While using IBM N series storage, local NTFS storage can be configured using FCP or iSCSI, and the storage can be easily managed by the IBM N series with SnapDrive solution. Using SnapDrive and IBM N series with FlexVol, the storage growth can be scaled according to the needs of data growth. IBM N series storage provides RAID protection and allows creating an instant backup and quick recovery using Snapshot and IBM System Storage N series with SnapRestore[®]. Both backup and recovery can be managed by SnapDrive. SnapDrive is integrated seamlessly with the MMC. This meets the requirement of SAN management for installing and configuring the database data files and transaction log files.

Installing File-System Archival Component

FSA is installed using the installation utility of Enterprise Vault software; uncheck the Enterprise Vault services and the administration console, leaving the file placeholder services, and follow instructions to add a new file server for archiving. See Figure 8.

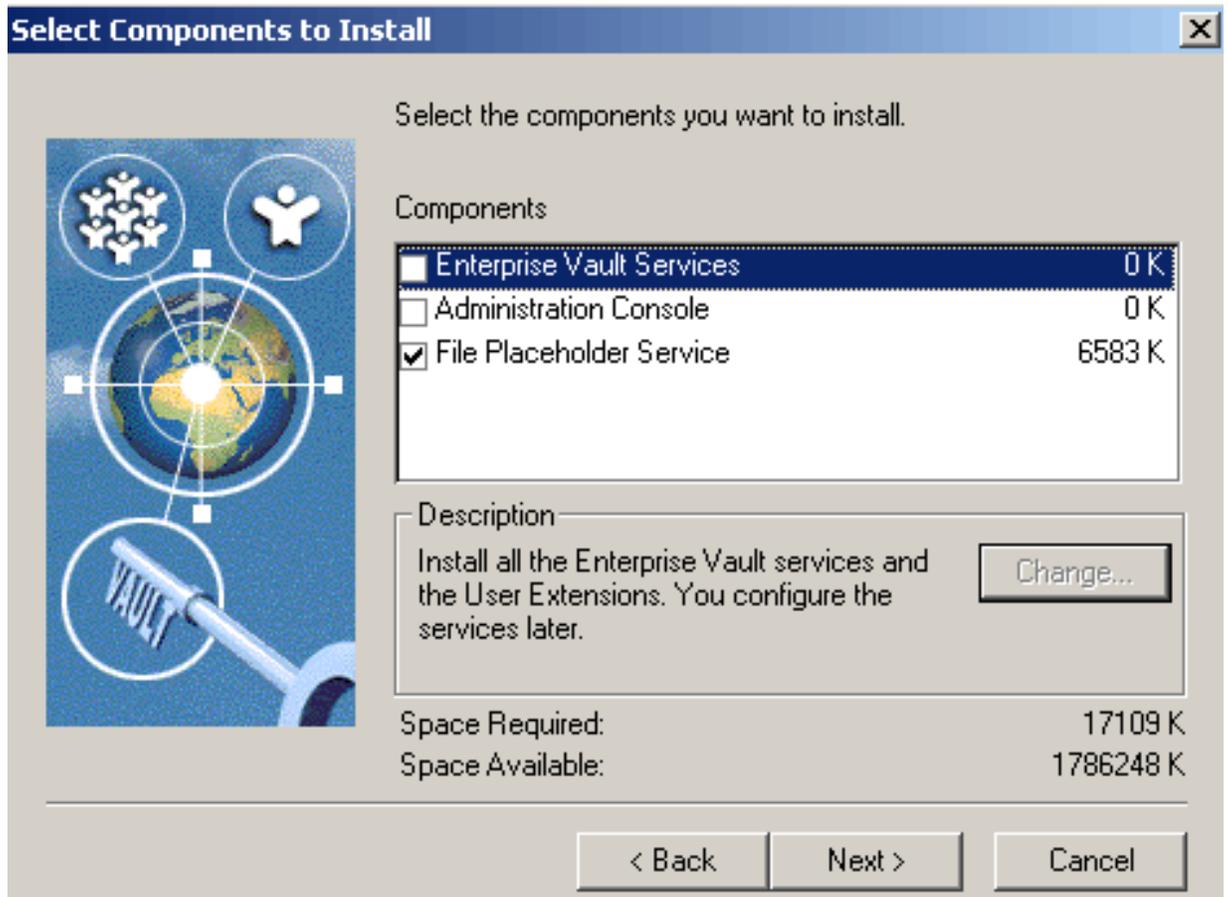


Figure 8. Installing file-system archival components.

Installing and configuring the FSA include adding the placeholder service to the target file server, adding a file server within the admin console, and creating the volume and folder policies. Before configuring FSA, verify that the Enterprise Vault properties are set properly. Configure the Enterprise Vault administrator user to have the appropriate permissions to do the following:

- Log on as a service
- Act as part of the operating system
- Debug programs.

Setting the locations for the vault directory database and transaction log is specified to store the database data files and transaction log files. Using the Enterprise Vault configuration wizard, create a new vault site. In our test setup, we created a vault site as shown in Figure 9.

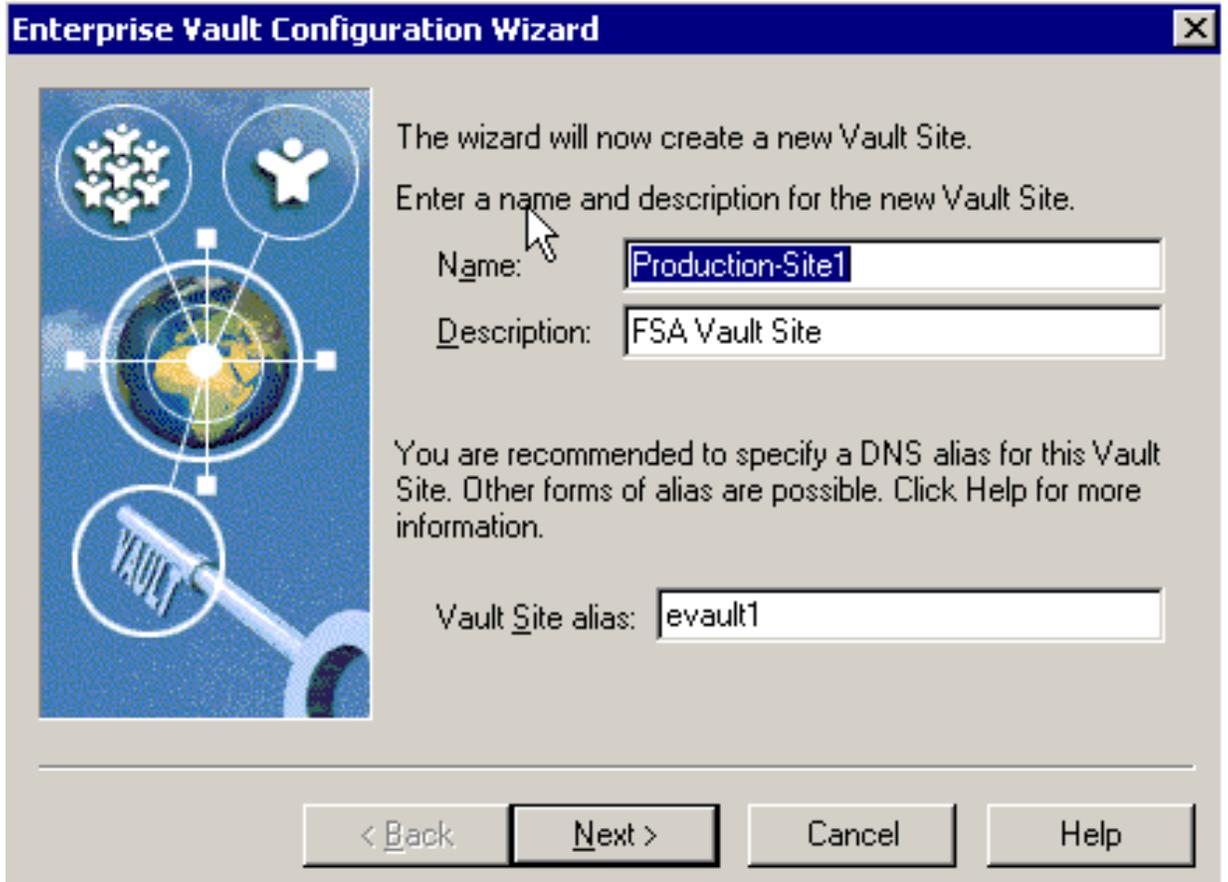


Figure 9. Configuring a new vault site.

While installing the placeholder service, verify that the correct properties are set, such as specifying the vault directory, directory, directory security, and permissions to execute scripts. In our test setup, we used the properties as shown in Figure 10.

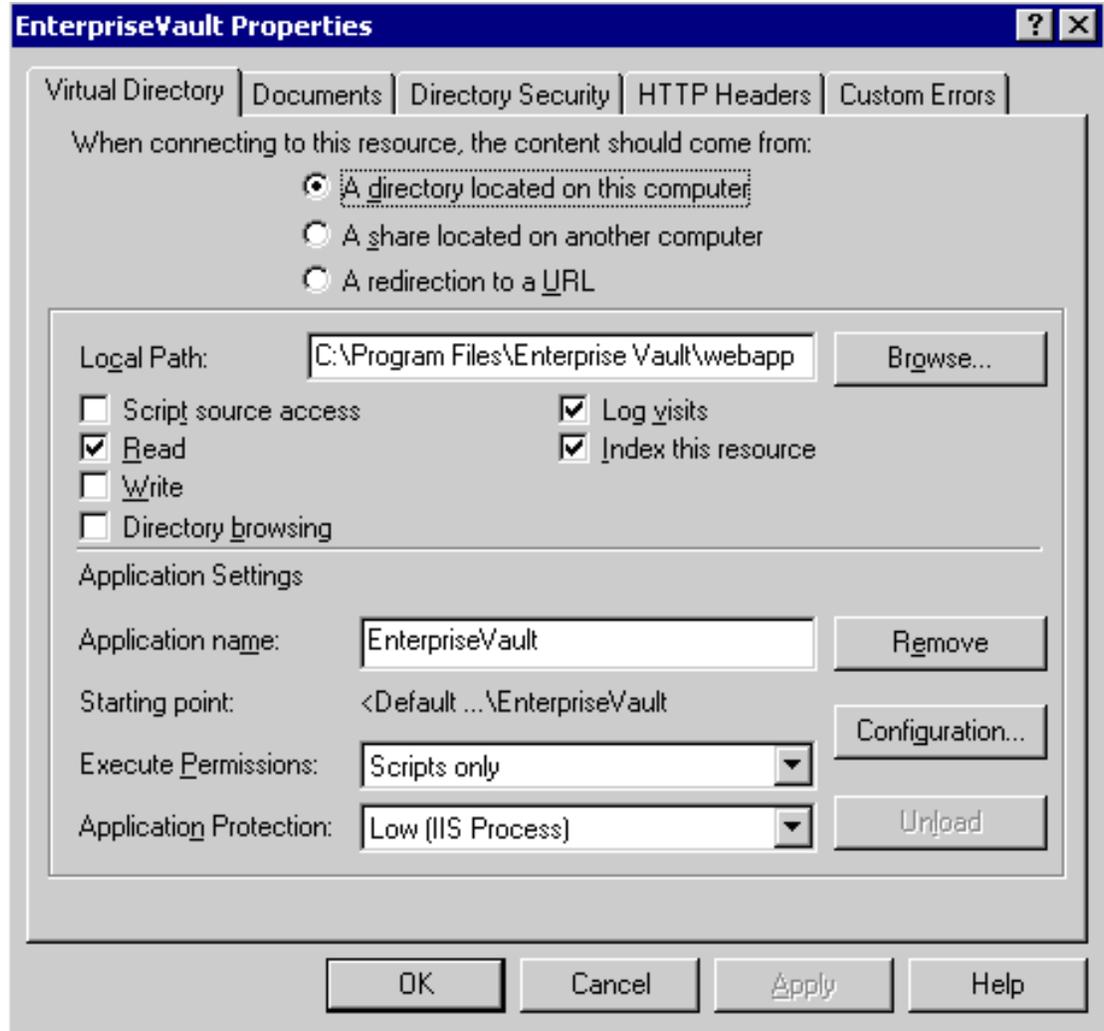


Figure 10. Enterprise Vault properties.

Enterprise Vault Configuration

In order to complete the FSA component installation, configure the Enterprise Vault server. All Enterprise Vault services must use a Windows account to log on. A vault service account is configured with an appropriate Windows user name and password. Typically, "vaultadmin" user will have necessary logon permissions to access the directory services. The service account will be configured to have user rights to do the following:

- Log on as a service
- Act as part of the operating system
- Debug programs.

This will allow you to set the necessary logon permissions and to install the FSA component on the Enterprise Vault server. See Figure 11.

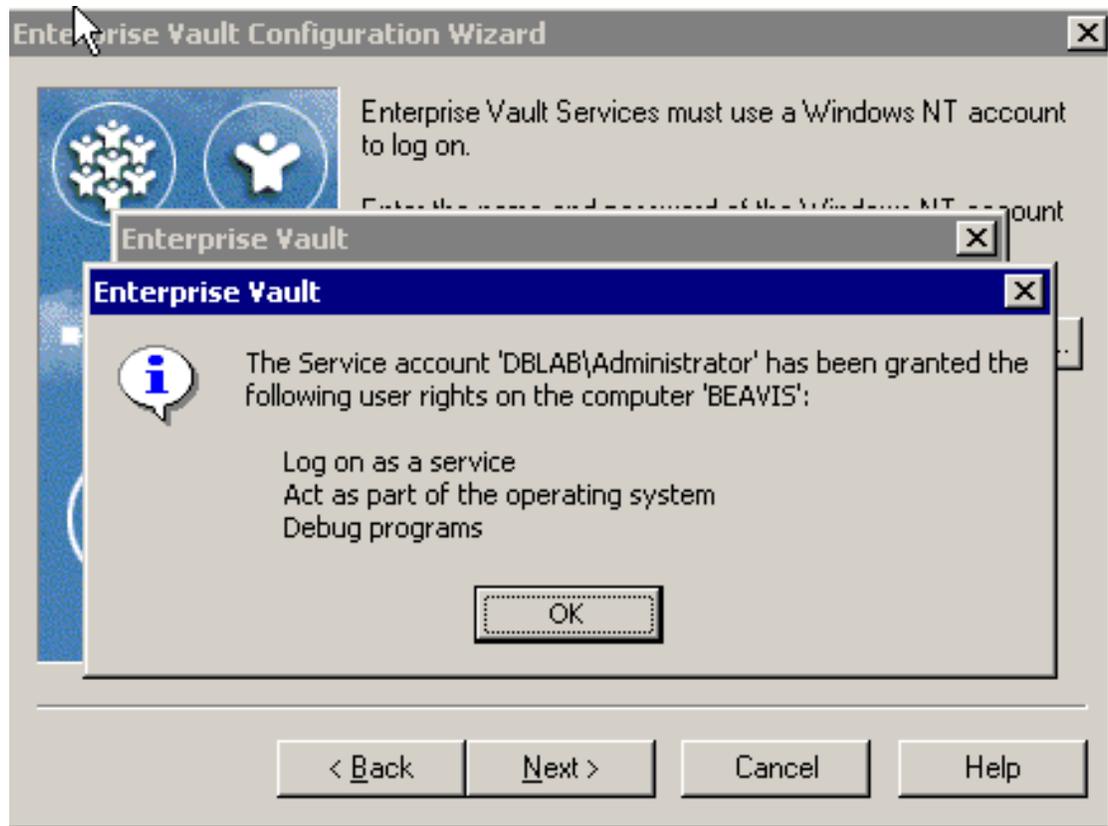


Figure 11. Enterprise Vault configuration.

Adding Vault Services

The Enterprise Vault configuration is continued by specifying the vault directory database and transaction log locations. In our test setup, we created a new vault site called Production-Site1 with the vault site alias "evault1." The configuration wizard obtains the vault directory computer and detects the software components available with Enterprise Vault, as shown in Figure 12.

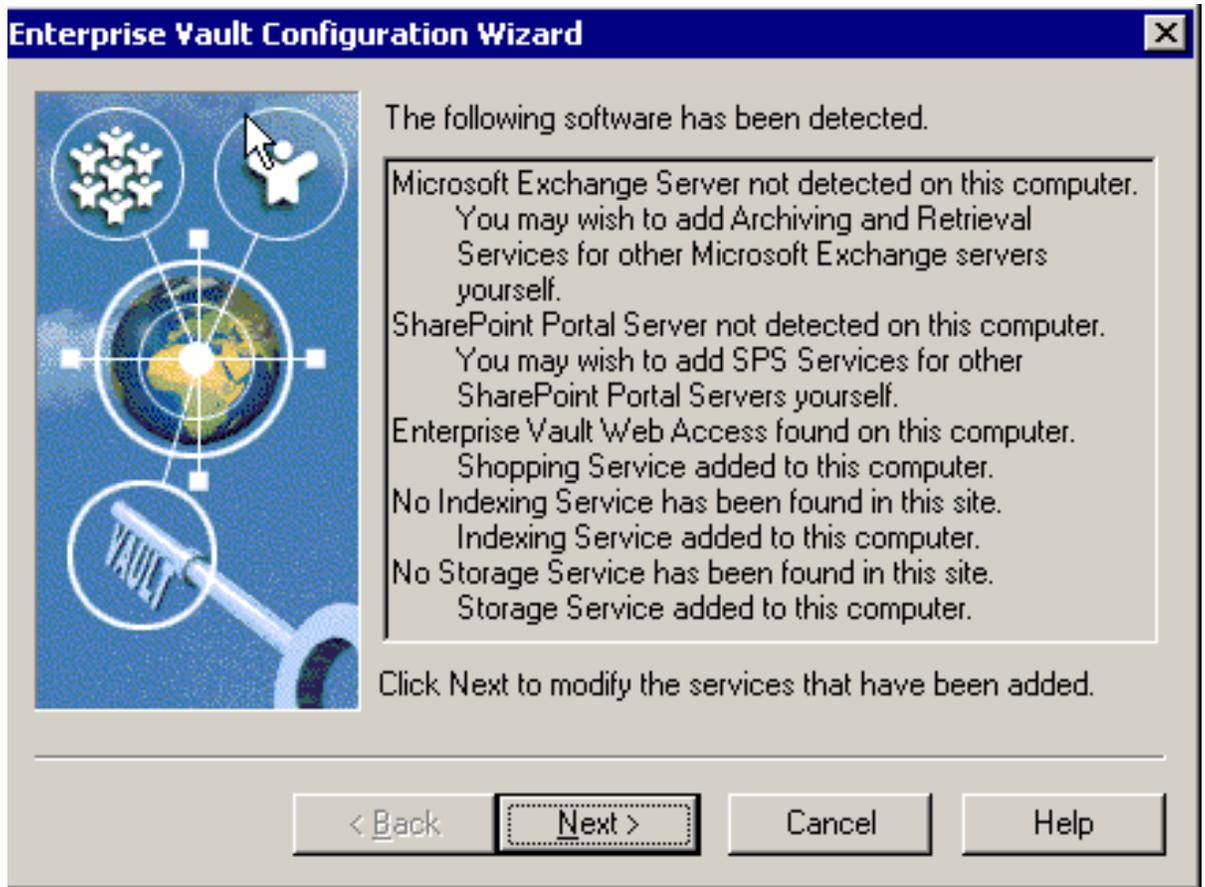


Figure 12. Service components installed and added.

In order to use FSA, Enterprise Vault has to be configured with the necessary services, such as indexing, shopping, and storage services. Once these services are configured, installed services will be displayed with their status. Start these services and verify that the services are started as shown in Figure 13.

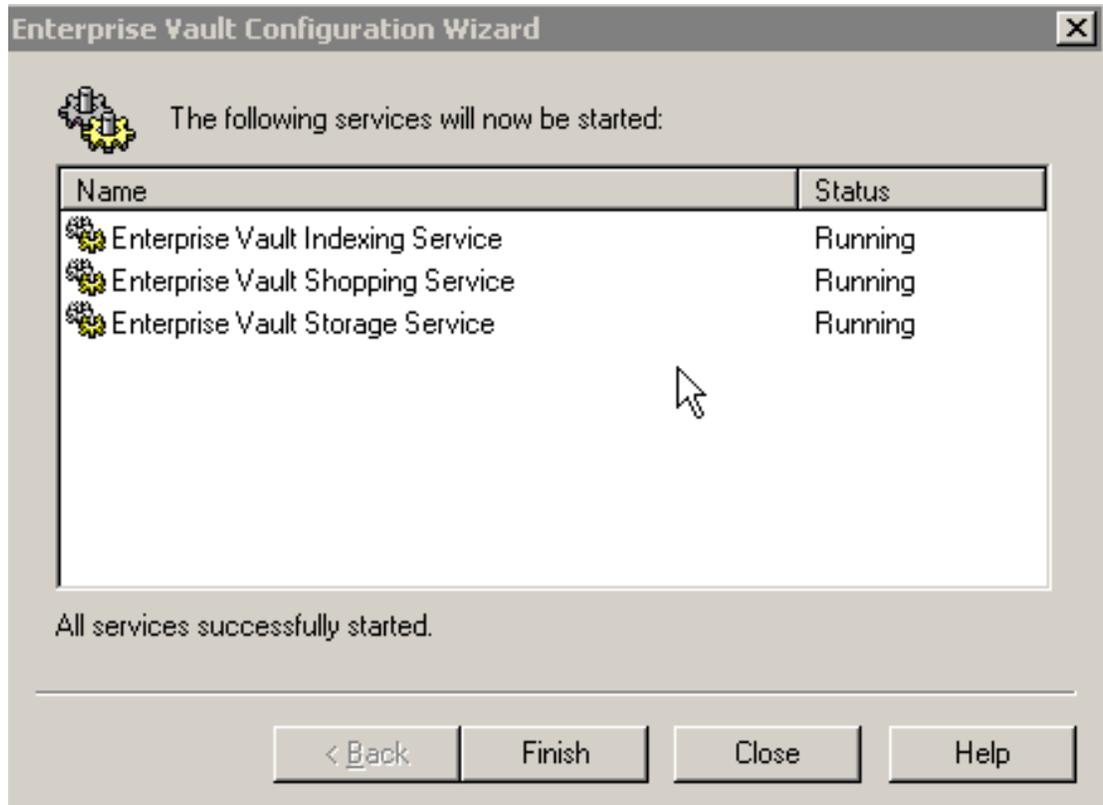


Figure 13. Enterprise Vault configuration wizard.

To complete FSA, create a new vault store to define the storage allocated to the partitions to archive the contents. Each vault store uses its own database to hold details of the archives within the vault store. Before creating archives, create a new partition on the vault store. It is important to specify whether to retain a safe copy once an item is archived or not. On our system, a new vault store and SQL Server database have been installed, as shown in Figure 14.



Figure 14. New vault store configuration.

Enterprise Vault can create a vault store partition on various storage devices such as a network share or IBM N series with SnapLock device, which provides non-erasable, non-rewritable WORM volume storage capability. Using Enterprise Vault and SnapLock, Enterprise Vault users and administrators can configure WORM-volume storage. In order to continue with the configuration, provide a location enabled with SnapLock for the new vault store partition, as shown in Figure 15.

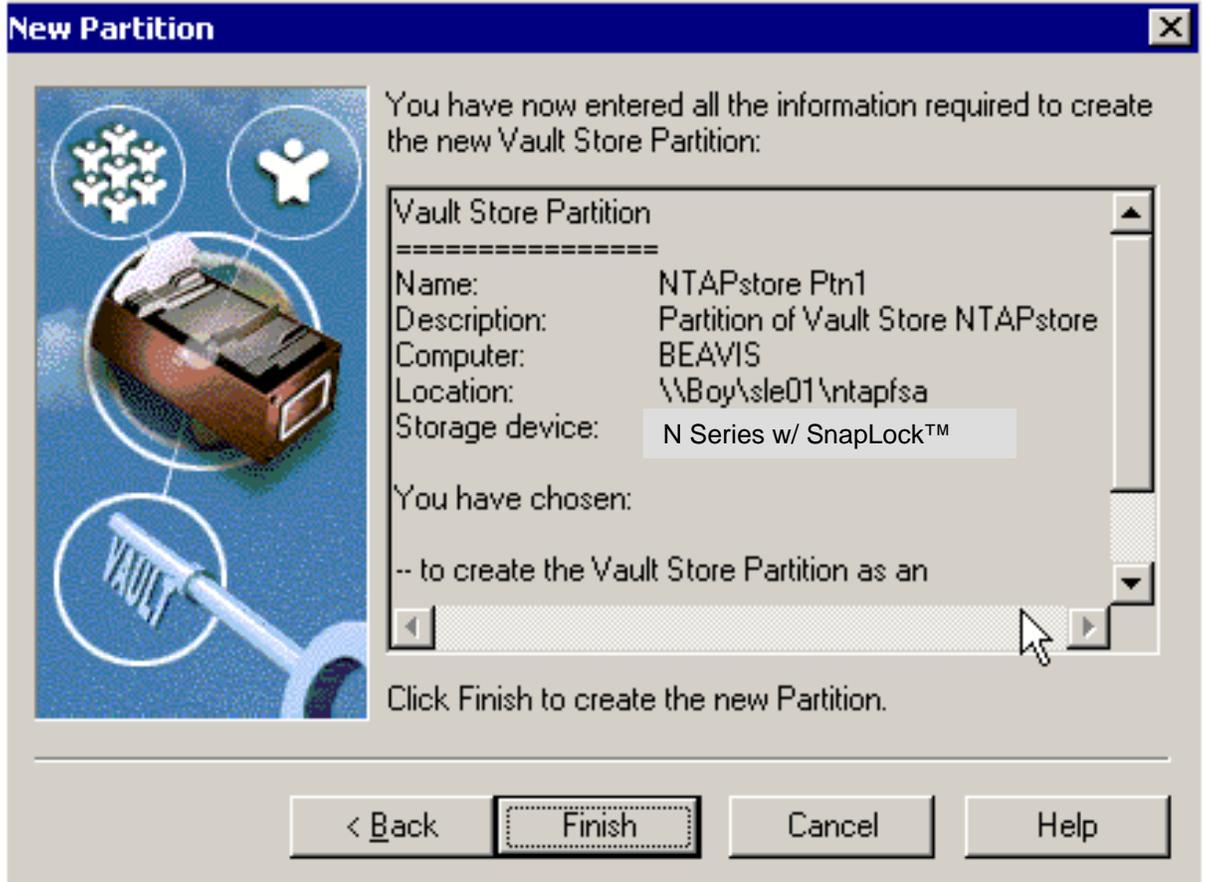


Figure 15. IBM N series SnapLock WORM volume storage config. path.

Adding a File Server

After installing the FSA component, we added the placeholder service to target file servers. The next task is to add the file server within the vault admin console and then create the volume and folder policies. In order to add a file server in the vault admin console, select the computer with storage services and the necessary path and verify that the new file server is added and enabled for use. Note that the new file server will have a default schedule for archiving, and you may change later to modify the default schedule. On our test setup a new file server, "Beavis", is added successfully. See Figure 16.

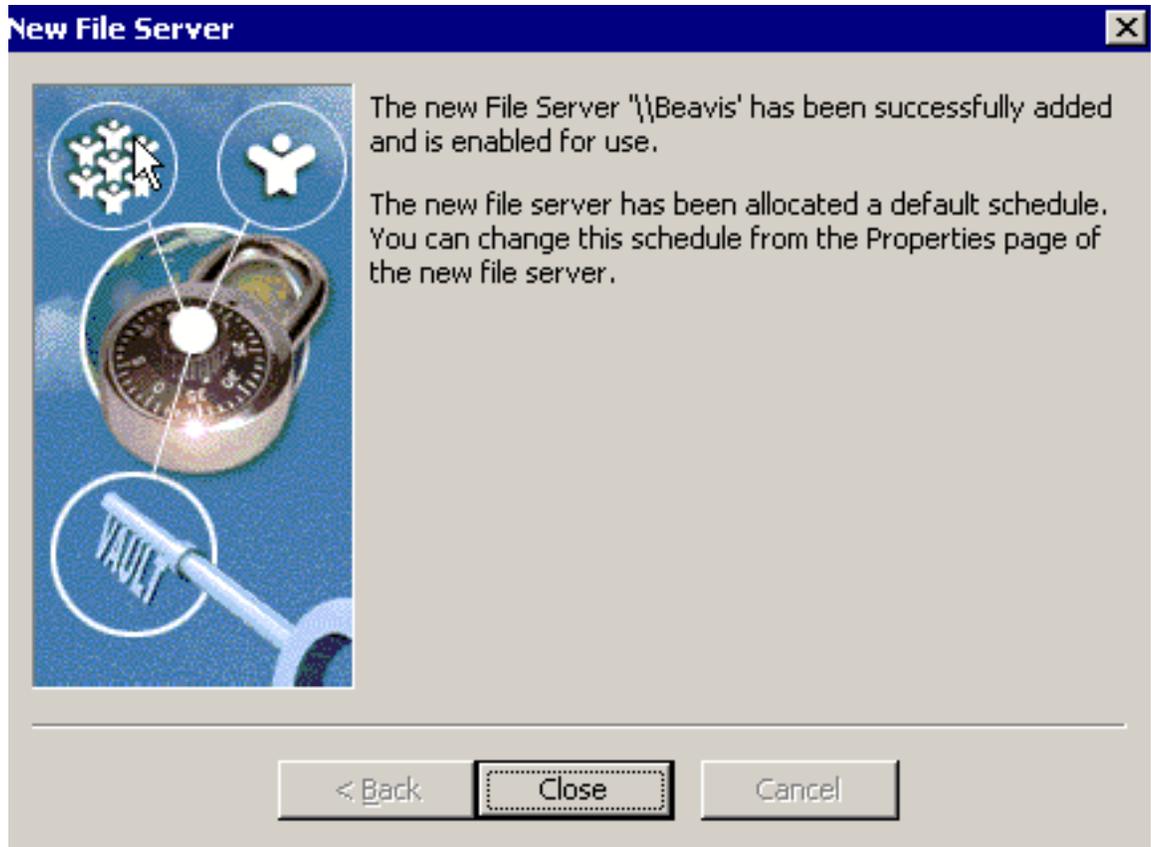


Figure 16. Adding a new file server.

Creating a New File Server Archiving Policy

FSA uses very flexible systems of policies and rules to control archiving. In order to provide greater flexibility to system administrators, different archiving policies may be enforced to different volumes or folders within a volume. However, each volume must have an archiving policy assigned to it. Folder policies can inherit or override the various settings from the overall volume policy. An example of a new policy is shown in Figure 17.



Figure 17. Creating a new policy.

The configuration includes to inform the Enterprise Vault server which file servers contain the volume and folders to be archived. The file server can be any computer on the network accessible to the Vault server. A single Enterprise Vault server can archive data from several file servers. However, each file server can be processed by one Enterprise server. Note that the configured Enterprise Vault server must be running the storage services.

The Vault administrator simply specifies the policies for the files to be archived and its source location, etc. The Enterprise Vault server will create the necessary required new processes. The placeholder shortcuts can only be used on devices hosted by a Windows server.

To add a new file server, open the Enterprise Vault Administration Console (EVAC) and expand until the File Servers container is visible and right click on the File Server and continue with creating a new file server.

Once the necessary File Servers are created, define the archiving policies for rules specifying the volume and folder policies as shown in Figure 17. Volume and folder policies include the rule the amount of free space limits on storage to start or stop archiving, retention category and other parameters. Policies define the method for selecting the files to archive and its method of archiving and handling of the original file after the archival process. Creating volume policies, folder policies and adding volumes can be created using the Enterprise Vault administration console. Figure 18 shows the File Servers configured on our test setup

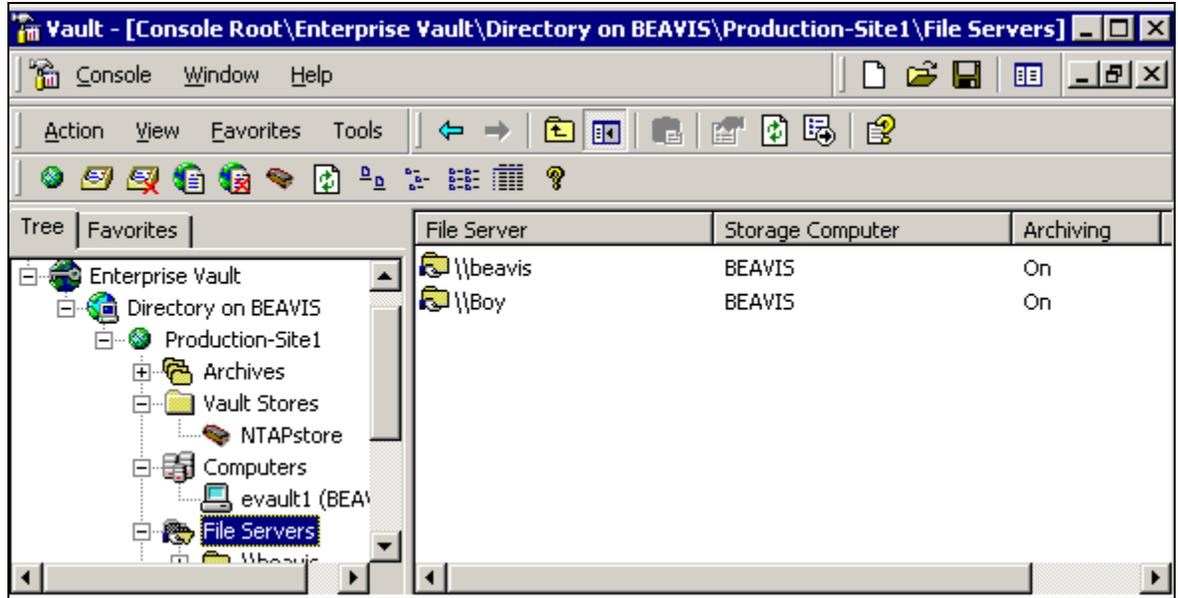


Figure 18. File server configured.

Policy Creation for Volumes and Folders

Before creating the volume and folder policies, it is important to know some useful tips such as the following:

- The archiving policy rules control exactly which files are to be archived and which must be excluded from archiving
- Be aware that a rule applies only when all the criteria are met
- Be sure not to apply too many rules by keeping them simple and error free
- Before archiving, perform an archive run in report mode
- Always *exclude* system files from being archived.

Scheduling File System Archiving

FSA is based on either site-level or file server-level schedules. File server-level scheduling is considered useful for staggering purposes. To schedule FSA, run the FSARunNow command. FSA processes each file server according to the schedule that is defined on a particular file server. Use the Enterprise Vault administration console to configure a schedule for a file server. See Figure 19.

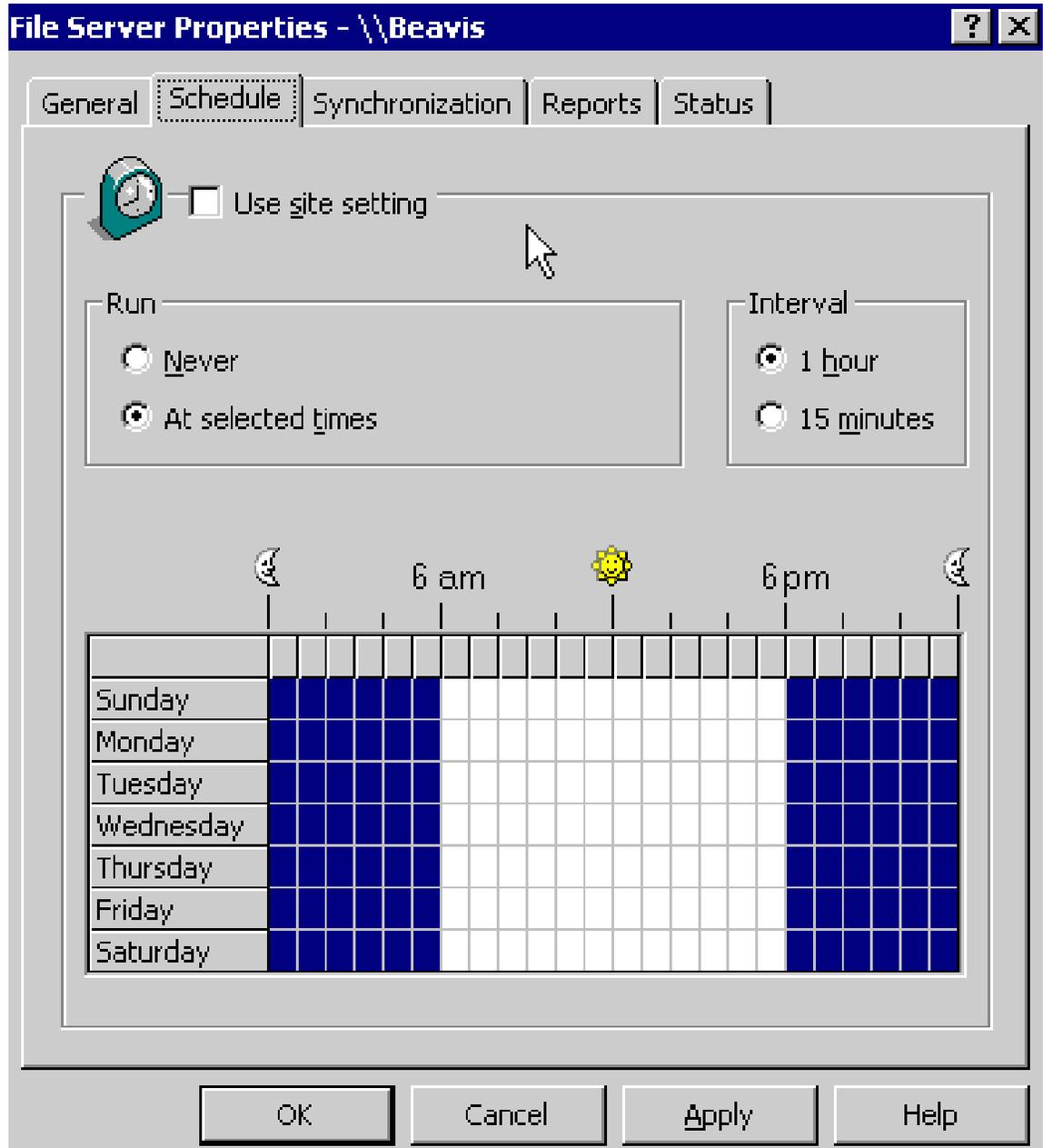


Figure 19. File server configuration schedule.



```
C:\Program Files\Enterprise Vault>archivepoints create
\\boy\sle01\ntapfsa\docs
Created archive point for folder: \\boy\sle01\ntapfsa\docs
C:\Program Files\Enterprise Vault>archivepoints find
\\boy\sle01\ntapfsa\docs
Listing Archive Points ...
    Archive Point : \\boy\sle01\ntapfsa\docs
C:\Program Files\Enterprise Vault>
```

Archive points are created using the site-level archiving defaults. This can be overridden using an XML template file for archiving name, description, owner, and indexing level.

FSA configuration files also enable you to create and edit to tune the file system archiving; the following scenarios can be handled:

- Archiving to a IBM N series with SnapLock volume
- Using HTML shortcuts
- Archiving items from recycle bins
- Modifying maximum archive file size.

The FSA configuration file can be created by renaming the sample configuration file EvFileArcSvr.exe.config, available at the Enterprise Vault installation directory, to EvFileArcSvr.exe.config.

This completes the installation and configuration of FSA in the Enterprise Vault server.



Summary

With Enterprise Vault, messages are extracted from Exchange Servers and stored. This storage has to be presented as an NTFS file system. Instead of placing messages in a database, messages are stored as standard files. This approach will offer several benefits. Some of these benefits include:

- No complex maintenance is required, as compared to databases
- A particular item can be recovered in case of a corruption, as opposed to an entire data set
- Without affecting performance, more data can be stored
- Multiple Exchange Servers can be handled by one or a few Enterprise Vault servers.

The Enterprise Vault server has several drawbacks in terms of data availability and dependability. To access data in archived files or to access files, SQL Server must always be up and running. In case of database corruption, it has to be recovered from the backup copy, losing all the recently archived items (files). Enterprise Vault works only in Microsoft Exchange environments. Space savings due to FSA may be offset by creating a secondary copy in the form of HTML. The data replication could take a significant amount of time and resources. Creation of a second copy of an HTML file after the file is archived reduces the space savings from archiving and compressing. Restoring the corrupted database could be disastrous in an enterprise environment.

The disadvantages just described can be easily addressed by exploiting the advantages of IBM N series storage solutions. The Symantec Enterprise Vault and IBM N series product integration design take advantages of both Enterprise Vault and the IBM N series storage solution to offer an efficient and highly available data solution.

IBM N series and Symantec are committed to providing Enterprise Vault users with superior solutions designed to meet their business needs. IBM N series filers and data management solutions ensure Enterprise Vault data is protected and available 24x7. With IBM N series, you get solutions that are easy to use, deploy, and manage, with high availability and exceptional performance at the lowest total cost of ownership in the industry.

IBM System Storage N series with SnapManager[®] for Microsoft SQL Server is a complete data management solution that provides backup and restore features using Snapshot technology. By reducing backup and restore times, minimizing application outages, and consolidating database storage, SMSQL delivers a cost-effective solution for managing critical SQL Server databases.

In conclusion, the recommendations made in this paper are intended to be an overview of best practices for *most environments*. This paper should be used as a set of guidelines when designing and deploying Symantec Enterprise Vault. To ensure a supported and stable environment, familiarize yourself with the resources provided in this paper and involve an Exchange specialist if necessary.

Caveats

All possible combinations of hardware platforms and storage architecture and software options have not been tested. If you use a different Windows Server OS or a different version of Enterprise Vault, then significant differences in your configurations could exist that may alter the procedures necessary to achieve the set objectives outlined in this document.



Trademarks and special notices

© International Business Machines 1994-2007. IBM, the IBM logo, System Storage, and other referenced IBM products and services are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. All rights reserved.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Network Appliance, the Network Appliance logo, Data ONTAP, FilerView, Snapshot and WAFL are trademarks or registered trademarks of Network Appliance, Inc., in the U.S. and other countries.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

Information is provided "AS IS" without warranty of any kind.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.