Symantec Enterprise Vault and EMC Centera

Applied Best Practices

EMC
Corporate Headquarters
Hopkinton, MA 01748-9103
1-508-435-1000
www.EMC.com

Published December, 2009

Symantec Enterprise Vault and EMC Centera Applied Best Practices

P/N H6790

Contents

About this Document 9
Introduction 9
Audience 10
Related documents 11
Symantec Enterprise Vault 8 manuals 11
EMC documentation 12
Chapter 1 Planning 15
Solution overview 16
Building block 18
Sizing considerations 19
Exchange 20
SQL Server 22
Enterprise Vault 26
EMC Centera 29
Network 30
Infrastructure / business continuity considerations 30
Enterprise Vault backup 31
Enterprise Vault clustering 32
Centera Replication 33
Centera Restore/tape out 35
Chapter 2 Implementation 37
Installation requirements and configuration 38
Enterprise Vault 38
EMC Centera 45
Compliance Mode 52
Installation process 54
Licensing 54
Installing Enterprise Vault 55
Post-installation tasks 56
Customizing security for the Web Access application 56
Customizing security on the client computers 57
Configuring Enterprise Vault 57
Setting up archiving 59

Custom configuration of Enterprise Vault and EMC Centera best practices 75
Enterprise Vault fine tuning 79
Installation validation 80
Enterprise Vault 80
EMC Centera 81
Chapter 3 Operations 83
EV updates and upgrades 84
Day-to-day management tasks 84
EMC Centera upgrades 84
Added capacity 85
CentraStar upgrades 85
EMC Centera migrations 85
Disaster recovery 86
Enterprise Vault only disaster recovery 86
EMC Centera primary only disaster recovery 87
EMC Centera Backup 88
Monitoring 88
High availability 89
Chapter 4 Performance Expectations 91
Enterprise Vault service level agreement (SLA) 92
Performance benchmarks 92
Ingest 93
Retrieval 94
Storage Expiry 94
Re-indexing 94
Accelerator applications 94
Stress tests 95
Cooperative Support 95
Chapter 5 Support Model 97
Symantec Enterprise Vault considerations 98
Enterprise Vault logging with DTRACE 98
EMC Centera considerations 99
SDK logging 99
Application Registration 100
Audit logging (syslog) 100
Alerts 101
Chapter 6 Conclusion 103
Appendix A Glossary 105
Terminology 106

Figures

Figure 1 E-mail archiving building block 19
Figure 2 Replication failover 34
Figure 3 Bidirectional replication with hot standby 35
Figure 4 Setting up removal of safety copies during vault store creation 63
Figure 5 Setting up removal of safety copies for an existing Vault Store 64
Figure 6 Selecting Centera as the storage type 65
Figure 7 Entering the EMC Centera access node (AN) IP addresses 66
Figure 8 Add Address (DNS Name optional) 66
Figure 9 Enabling Centera SiS on the Enterprise Vault partition 67
Figure 10 Setting up single instancing (sharing) 68
Figure 11 Setting retention policies from Enterprise Vault to Centera 69
Figure 12 Defining Centera Replication's partition scan intervals 70
Figure 13 Setting up Centera collections 71
Figure 14 Enabling Centera collections on an Existing Vault Store partition 72
Figure 15 Associating Enterprise Vault Retention Categories and EMC Centera Retention Classes 75
Figure 16 Establishing the EMC Centera connection string 76
Figure 17 Number of threads 78
Figure 18 Provisioning Group Properties 79

Tables

Table 1 Recommended number of SQL Server servers 24
Table 2 Recommended database storage allocation 25
Table 3 Supported mailboxes for mailbox archiving based on EV server configuration 27
Table 4 Index storage requirements 28
Table 5 Additional storage requirements per Enterprise Vault Server 29
Table 6 Minimum specification for an Enterprise Vault server 41
Table 7 Minimum specification for SQL Server 43
Table 8 Recommended EMC Centera configuration 45
Table 9 Enterprise Vault components – Summary of recommended values for EMC Centera devices 62
Table 10 Recommended default retention category for EMC Centera 74
Table 11 Expected Enterprise Vault ingest rates for physical CPUs 92
Table 12 Performance and stress test results 92
Table 13 Performance test components configuration 93
Table 14 Required SDK environment variables 100

About this Document

Information organizations are faced with an environment that is becoming increasingly subject to external and internal regulations and supervision. The scope and number of requirements and regulations facing businesses today are increasing – along with the costs of ensuring they are complied with.

Even though the cost of compliance can be considerable, the legal and goodwill risks associated with noncompliance can be even higher.

Proper management of unstructured data growth, primarily driven by the use of e-mail as the predominant form of enterprise communications, imposes a greater challenge onto today's organizations, which are determined to provide their customers – internal and external – with exceptional service and a satisfying experience.

Enterprise Vault is a Windows application that enables companies to automatically apply corporate, legal, and IT policies around the storage, transfer, retention, and disposition of unstructured data such as e-mail, IM, files, or SharePoint documents. This translates into reduced and simplified infrastructure management costs, and controlled data exposure and protection while ensuring a centralized and indexed archive that can be searched on demand.

EMC Centera® is a simple, affordable, power-efficient, and secure repository purpose built for information archiving to keep static and infrequently changing digital information available online for immediate access. EMC Centera enhances business value by preserving original content and ensuring complete, reliable integrity for the life of the archived information.

EMC Centera's self-healing/self-management/self-configuration capabilities, in addition to nondisruptive repair and phone-home features, enable customers to efficiently manage more archive content at a lower cost.

Supported by unmatched retention and disposition management capabilities, EMC Centera successfully addresses the most rigorous compliance regulations, while minimizing the total cost of ownership. EMC Centera enforces organizational and application policies for information retention and disposition intrinsic in storage—and by doing so completes the information chain of custody. It ensures corporate accountability and reduces the costs of legal discovery and litigation support—and it's easy to manage.

From application through storage, EMC Centera's unique functionality, in combination with Symantec Enterprise Vault's flexibility, creates an intelligent, complete chain of information custody. This level of information authenticity and the ability to ensure that an organization's retention and disposition policies are enforced make EMC Centera and Symantec Enterprise Vault the optimal yet simple e-mail and file system archive solution for businesses and organizations.

Introduction

This white paper addresses the challenges of deploying an e-mail archiving solution by looking at the several phases of any project life cycle, from planning and architecting to installing, configuring, and fine-tuning of the system, in addition to covering maintenance activities after

successful deployment, in order to provide guidelines and recommendations based on lessons learned through several years of successful partnership and tight technical integration between Symantec and EMC.

The paper is not intended to replace the documentation that either Symantec or EMC provide for Symantec Enterprise Vault and EMC Centera, but rather to serve as a reference and complement to the already existing documentation, and to provide stakeholders with a holistic, unified perspective on the topic of e-mail archiving as it pertains to the joint offering.

In this context, the paper discusses not only the specifics of installing and configuring Centera and Enterprise Vault, but also the additional know-how EMC has on deploying Microsoft Exchange and Microsoft SQL Server as the e-mail and database engines on the appropriate storage tiers for the enterprise. Similarly, an overview of the expectations around performance and available support mechanisms is provided.

Although Centera configuration is minimal and is performed as part of the installation process, detailed information on each topic is provided in order to minimize ambiguity, and misinterpretation, and to provide a greater understanding of the benefits each feature provides to the organization.

As it is the role of the application to exploit EMC Centera capabilities, careful attention to Enterprise Vault's archiving configuration is required, and dedicated in the paper.

The ultimate goal for this document is to trigger an open dialogue between all parties involved and to clarify the most common and critical project/product areas in order to successful identify, plan, implement, monitor, and maintain the several component parts of the solution.

Audience

This white paper is intended for:

- Customers, including IT planners, storage architects, and administrators
- EMC and Symantec technical staff and partners
- Field personnel who are tasked with implementing archiving solutions, using Symantec Enterprise Vault and EMC Centera, focused on e-mail archiving for Microsoft Exchange Server environments
- Engineering and product development groups
- Account personnel involved in pre-sales activities

This white paper assumes the reader has a basic knowledge and understanding of Microsoft Windows, Microsoft Exchange Server 2007, Microsoft SQL Server, Symantec Enterprise Vault, and EMC Centera.

Related documents

Symantec Enterprise Vault 8 manuals

Core guides

TechNote Number	Document Title/Link	File Name
317272	Symantec Enterprise Vault 8.0 - Administrators Guide	Administrators_Guide.pdf Administrators_Guide.chm
312327	Enterprise Vault 8.0 and later - Backing up Enterprise Vault	Enterprise Vault 8.0 and later - Backing up Enterprise Vault (PDF)
317273	Symantec Enterprise Vault 8.0 - Introduction and Planning Guide	Introduction_and_Planning.pdf Introduction_and_Planning.chm
312319	Symantec Enterprise Vault 8.0 Performance Guide	http://seer.entsupport.symantec.com/docs/312319.htm
317276	Setting_up_Domino_Server_Archiving	Setting_up_Domino_Server_Archiving.pdf Setting_up_Domino_Server_Archiving.chm
317278	Setting up Exchange Server Archiving	Setting_up_Exchange_Server_Archiving.pdf Setting_up_Exchange_Server_Archiving.chm
317287	Setting_up_File_System_Archiving	Setting_up_File_System_Archiving.pdf Setting_up_File_System_Archiving.chm
317288	Setting_up_SharePoint_Server_Archiving	Setting_up_SharePoint_Server_Archiving.pdf Setting_up_SharePoint_Server_Archiving.chm
317290	Setting_up_SMTP_Archiving	Setting_up_SMTP_Archiving.pdf Setting_up_SMTP_Archiving.chm
317274	Registry Guide	Registry_Guide.pdf Registry_Guide.chm
317275	Symantec Enterprise Vault 8.0 - Utilities Guide	Symantec Enterprise Vault 8.0 - Utilities Guide.chm
276547	Symantec Enterprise Vault (tm) 6.0, 7.0, 2007 and 8.0 Compatibility List	Veritas Enterprise Vault (tm) 5.0, 6.0 and Symantec Enterprise Vault 7.0, 8.0 Supported Products list (includes Partner products).pdf
N/A	Backing up Enterprise Vault in a clustered environment – **For internal distribution only**	Backing up Enterprise Vault in a clustered environment.pdf
N/A	Symantec Enterprise Vault 7.0 - Sizing an Enterprise Vault Solution for Exchange Archiving – Technical white paper – **For internal distribution only**	Symantec Enterprise Vault 7.0 - Sizing an Enterprise Vault Solution for Exchange Archiving – Technical white paper.pdf

Install/Upgrade

TechNote Number	Document Title/Link	File Name
317303	Symantec Enterprise Vault 8.0 - Deployment Scanner Guide	Deployment_Scanner.pdf Deployment_Scanner.chm
317266	Symantec Enterprise Vault 8.0 - Installing and Configuring Guide	Installing_and_Configuring.pdf Installing_and_Configuring.chm
298949	Release Notes 8.0	Symantec Enterprise Vault 8.0 Release Information

Compliance Accelerator and Discovery Accelerator 8 manuals

TechNote Number	Document Title/Link	File Name
317294	Compliance Accelerator - Installing and Configuring Guide	Installing_and_Configuring_Compliance_Accelerator.pdf
317300	Discovery Accelerator - Installing and Configuring Guide	Installing_and_Configuring_Discovery_Accelerator.pdf
317296	Compliance Accelerator - Reviewer Guide	CA_Reviewer_Guide.doc
317298	Discovery Accelerator - Reviewer Guide	DA_Reviewer_Guide.doc
326055	Discovery Accelerator - Best Practices for Implementation	Symantec Discovery Accelerator 8.0 - Best Practices for Implementation

EMC documentation

Note: To access EMC documents that are published on Powerlink, a user must have a Powerlink account or contact an EMC representative.

EMC Centera

Part Number	Document Link	File Name
H4498	EMC CentraStar 4.1 Network Segmentation - A Detailed Review	EMC CentraStar 4.0 - Network Segmentation A Detailed Review
069001127 REV A07	Centera V2.1 Programmer's Guide	SDK Programmers Guide
069001185 REV A08	EMC Centera SDK V3.2 API Reference Guide	EMC Centera SDK V3.2 API Reference Guide
H4496	EMC Centera Virtual pools Introduction and Principles of Operation - A Detailed Review	EMC Centera Virtual Pools Introduction and Principles of Operation - A Detailed Review
H5553	EMC Centera Replication - A Detailed Review	EMC Centera Replication - A Detailed Review
H5735.3	EMC Centera CentraStar/SDK Compatibility with Centera ISV Applications - A Detailed Review	EMC Centera CentraStar / SDK Compatibility with Centera ISV Applications - A Detailed Review
H5792	EMC Centera Network Segmentation Best Practices Planning	EMC CentraStar 4.0 - Network Segmentation Best Practices
H5874	Updating EMC Centera Access Node IP - Best Practices Planning	Updating EMC Centera Access Node IP - Best Practices Planning
N/A	**For internal distribution only**	Symantec Partner Kit
N/A	Centera CentraStar and SDK Release and Interoperability Matrix	Centera CentraStar SDK Interoperability Matrix
H4497	http://powerlink.emc.com/km/live1/en_US/Offering_Technical/White_Paper/H4497-emc-centera-cmplnce-mdls-gvrnce-edtn-cmplnce-edtn-plus-wp.pdf	EMC Centera Compliance Models - A Detailed Review

Microsoft Exchange

	Document Link	File Name
H4060.1	http://powerlink.emc.com/km/live1/en_US/Offering_Technical/White_Paper/H4060-emc-clariion-stor-sol-ms-exch-07-wp.pdf?	EMC CLARiiON Storage Solutions: Microsoft Exchange 2007 – Best Practices Planning
N/A	http://www.emc.com/collateral/hardware/technical-documentation/esrp-tech-docs.htm	Microsoft Exchange Solution Reviewed Program (ESRP)

SQL Server

	Document Link	File Name
H4060.1	http://powerlink.emc.com/km/live1/en_US/Offering_Technical/White_Paper/H4060-emc-clariion-stor-sol-ms-exch-07-wp.pdf?	EMC CLARiiON Storage Solutions: Microsoft Exchange 2007 – Best Practices Planning
N/A	http://www.emc.com/collateral/hardware/technical-documentation/esrp-tech-docs.htm	Microsoft Exchange Solution Reviewed Program (ESRP)

Third-party documentation

Part Number	Document Link	File Name
N/A	http://www.enterprisestrategygroup.com/ViewSecureDocument.asp?ReportID=827&ReportType=Free&ReportField=Attachment1	ESG Lab Review EMC Centera Email Archiving Jun 07 FINAL[1].pdf

For Microsoft SQL Server information, refer to the Microsoft websites, including http://www.microsoft.com/sql/, or more directly to the documentation page for the various versions at: http://msdn.microsoft.com/sql/sqlref/docs/default.aspx.

Microsoft SQL Server Books On Line documentation provides extensive coverage of features and functions and may be installed through the SQL Server installation process, independently of the SQL Server database engine. Updated versions of the Books On Line documentation are available for free download from the Microsoft SQL Server website: http://www.microsoft.com/sql

Note: Links listed are functional as of this document's publication date. Over time, the location of reference material may change. A navigation description is included to assist in those cases. If the user has difficulty locating referenced material, contact your EMC or Symantec representative.

Chapter 1 Planning

This chapter presents these topics:

Solution overview 16
Building block 18
Sizing considerations 19
Infrastructure / business continuity considerations 30

Solution overview

Enterprise Vault provides the framework for organizations to automatically store messaging and file system data from Microsoft Exchange, SharePoint Portal Server, Lotus Domino, SMTP, IM, file server environments, and other collaborative environments into centralized archives.

In particular, automatic e-mail archiving and archive management provide the following important benefits:

- It ensures that messages and documents are retained for the period of time required by compliance regulations or company policy.
- The size of Exchange Server mailboxes and public folders and Domino mail files is easily controlled without the loss of data (.PST or .MSG files).
- Primary disk space usage is reduced.

According to scheduled times, Enterprise Vault archiving processes check target servers for items to archive based on predefined policies. Items meeting the archival criteria are then stored in the archive. Archived items are indexed by Enterprise Vault to enable fast searching and retrieval, based on the level of indexing required by the organization as set up by the administrator.

A retention category, which defines how long an item must be kept, is automatically assigned to each item at archiving time. The administrator can define different retention categories for different data types or subsets of data.

EMC Centera® receives archived content from Enterprise Vault and stores this content as individual objects, which are uniquely identified by content addresses (CAs) derived by running a hashing algorithm over the archived content. Archived content is then accessed by Enterprise Vault using these CAs rather than by name or path as with traditional stores.

If an attempt is made to store the same file, regardless of what application it comes from again, EMC Centera will generate exactly the same CA and will detect that there is no need to store this redundant copy. For disaster recovery purposes, data is automatically replicated to the target or replica EMC Centera cluster in an asynchronous fashion.

Once Enterprise Vault verifies that the data is replicated, the original items can be replaced with shortcuts to the archived copy in order to release primary storage on user computers and target application servers. From the end user's point of view, they will still be able to access the items as before.

As Enterprise Vault monitors the archives, it can then delete items when the retention period expires. Under Centera compliance mode, any application attempt to delete data still under retention will be denied and an event log will be created

In order to better understand the data flow between Enterprise Vault and Centera, workflows for ingest, retrieval, and delete operations are discussed next.

The following is the workflow for content ingests:

Component	Steps
Enterprise Vault	1. Archiving Services pick up items from target stores (such as Exchange mailboxes, public folders, SharePoint workspaces) and queues them to the Storage Service. 2. Storage Service stores them and initiates Indexing Service. 3. Indexing Service indexes selected items' attributes. 4. Items are compressed. 5. Collections are built, if applicable. 6. Items are sent to EMC Centera for archiving (when Enterprise Vault is storing shareable data it sets EMC Centera's FP_OPTION_CLIENT_CALCID so that data already on the EMC Centera is not sent again)
EMC Centera	7. Content Address is calculated. 8. Single instancing is checked. 9. Item is stored based on storage policies. 10. Item is put in the replication queue.
Enterprise Vault	11. Archiving Services verifies the item exists in replica (scheduled / configurable). 12. A shortcut to the item is created once the replica copy is detected. 13. Item is removed from the target store (stubbed out).

The following is the workflow for content retrieval:

Component	Steps
Enterprise Vault	1. End user requests item from the information source (such as Exchange mailboxes, public folders, SharePoint workspaces). 2. Storage Service uses SavesetId to look up an item in the SQL database, and to get the Clip-Id. 3. The EMC Centera content address (Clip-id) is determined from database lookup.
EMC Centera	4. Content is read (Partial Read is performed per collection items).
Enterprise Vault	5. Item is decompressed and re-assembled into a saveset object. 6. The saveset is validated against SavesetID and clip attributes. 7. The saveset object is made available for the end user.

The following is the workflow for content deletion:

Component	Steps
Enterprise Vault	1. Expiry process identifies item to be deleted based on retention policies (alternatively, end user can also request an item to be deleted). 2. Storage Service uses SavesetId to look up an item in the SQL database, and to get the Clip-Id. 3. The EMC Centera Content address (Clip-id) is determined from database lookup.
EMC Centera (primary)	4. Content is read on the primary cluster (Partial Read is performed per collection items).
Enterprise Vault	5. The item is validated against SavesetID and clip attributes.
EMC Centera (replica)	6. The item is deleted (if collections are used, the counter is decreased – if the counter reaches zero (0), collection is deleted).
EMC Centera (primary)	7. The item is deleted (if collections are used, the counter is decreased – if the counter reaches zero (0), collection is deleted).
Enterprise Vault	8. Indexing Service deletes the entry from the AltaVista index file 9. Storage Service deletes Saveset and SavesetStore records (which includes the Clip-ID) from the SQL Server DB.

Building block

BEST PRACTICE: To facilitate the understanding and deployment of the core system components of the e-mail archival solution, a building block approach is recommended.

A building block is a repeatable unit of Enterprise Vault functionality, hosted on an application server, comprising a full set of Enterprise Vault Services that deliver archiving, indexing, storage, search, and retrieval functions and is both reliable and scalable.

This methodology makes it easy to extend Enterprise Vault capabilities by adding as many blocks/units as needed in order to support business requirements around capacity, performance, and resiliency. This is particularly important for disaster recovery purposes, as an extra building block could serve failover roles.

Figure 1 is a graphical representation of the typical Enterprise Vault and EMC Centera building block to support e-mail archiving for 7,500 mailboxes.

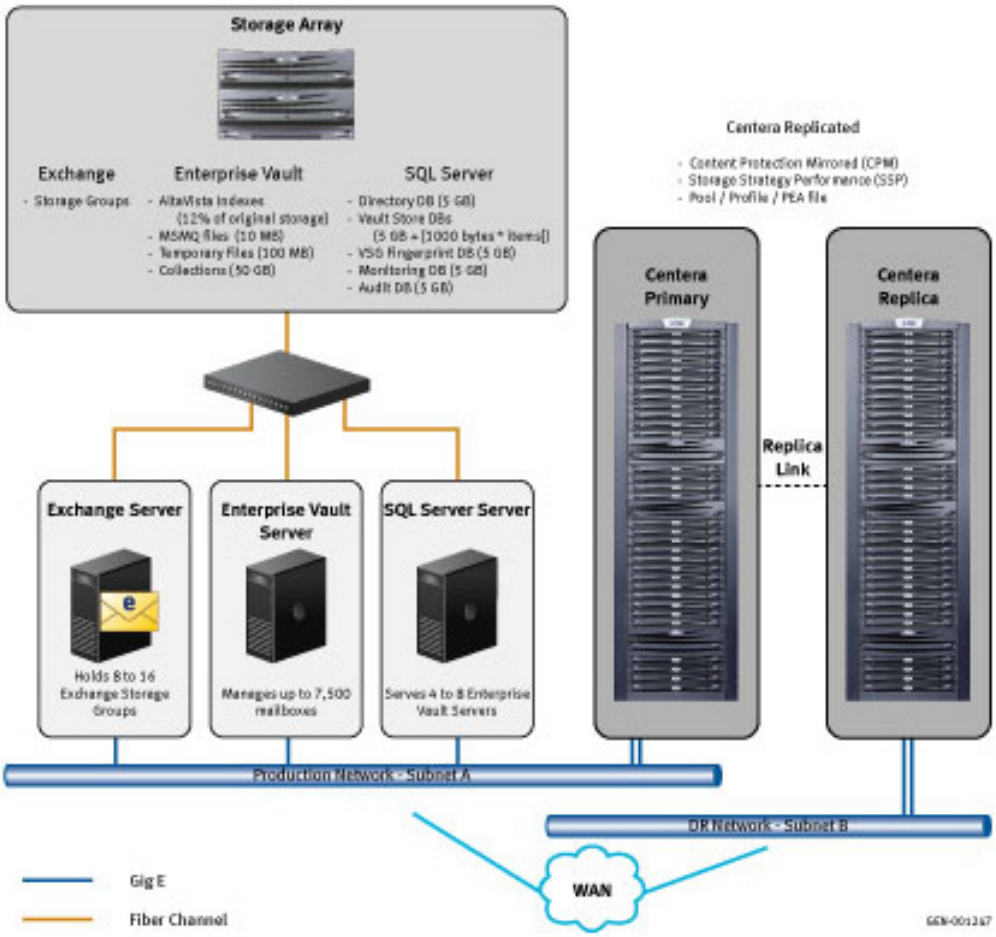

Figure 1 E-mail archiving building block

Sizing considerations

When sizing an e-mail archiving solution, there are several factors to consider such as:

- Mailbox data characterization
- Number of mailboxes
- Number of Exchange servers and their locality
- Type of archiving (journal or e-mail archiving)
- Windows timeframes to perform such activities as:
  - Network infrastructure

- Enterprise and divisional policies for archiving, backup, disaster recovery, and high availability

These environmental factors, unique to each organization, have a direct impact on the size of the solution components – mainly the configuration of the Enterprise Vault and SQL servers, and their associated storage requirements. This section will present an overview of each component and guidelines for determining its appropriate sizing and configuration to effectively and efficiently address the business requirements.

Exchange

Although Microsoft Exchange is presumed to be installed already, the following are some high-level considerations and recommendations for designing a successful Exchange storage system, based on benchmarks performed by EMC according to the Microsoft Exchange Solution Reviewed Program (ESRP) – Storage program.

The ESRP – Storage program was developed by Microsoft Corporation to provide a common storage testing framework for storage vendors to provide information on their storage solutions for Microsoft Exchange Server. For more details, visit http://www.microsoftstoragepartners.com.

User information

This information will help understand the I/O profile for a set of users, and what their storage requirements are.

Mailboxes

- Total number of mailboxes (regular e-mail and journal) and size limits
- Mailbox distribution across multiple Exchange servers/locations
- Daily amount of messages and size
- Anticipated growth over the next few years
- Retention period for *deleted items*

User activity

- Maximum I/O per user
- Average number of log files created per user per day
- Typical working day
- Peak activity periods and number of concurrent users at that time

Backup/Recovery

- Backup method
- Backup and maintenance schedules and window timeframes
- Recovery service level agreements (recovery point objective and recovery time objective)

- Disaster recovery distance from the primary site, network connection links, and available bandwidth

To assist gathering some of these metrics on an existing Exchange environment, the Windows System Monitor tool (PerfMon) can be used on a production server with a user load matching the target community. It is recommended that the run includes a 1-to-2 hour period during peak user activity. In particular, the following sample measurements could be useful:

1. IOPS per user: Use the System Monitor counter Physical Disk\Disk Transfers/sec on the Exchange database drive and the following formula:

   IOPS per User = (Disk Transfer/sec) / (Number of Users)

2. Read/Write Ratio: During the peak activity period, the ratio is measured on the database drive using the formula:

   Read/Write Ratio = (Disk Reads/sec) / (Disk Writes/sec)

3. I/O Latency: The read or write latency to the database or log drives can be measured using the counters:

   PhysicalDisk\Avg Disk sec/Read and Avg Disk sec/Write

Site-specific constraints

Any environmental requirements or restrictions where the messaging system is to be implemented must be taken into consideration such as:

- The use of a new or existing array and whether or not it is dedicated to Exchange or shared with other applications
- Type of drives in the array
- Number and location of Exchange servers
- Number of Exchange Storage Groups (ESGs) per server
- Network capacity

An ESG is the fundamental unit for layout planning; the following are considerations/recommendations relevant to ESGs:

- For backup purposes, the elements of an ESG should be treated together.
- For Exchange 2007, it is best to place a single database in each storage group.
- Although determining how many ESGs to have on a server is not as clear, EMC feels that eight or 16 ESGs per server provides a good balance between performance and manageability.
- Database size will vary based on customer requirements but should in general be kept under 210 GB in size.
- Under most circumstances, two LUNs should be allocated for each ESG: one for the database file and one for the transaction logs. These LUNs should not share spindles.

- For production Exchange database LUNs, there are two possible RAID type configurations that can be appropriate:
  - RAID 1/0 (mirrored), which offers the best performance with high protection, but only 50 percent of the RAID group capacity is usable. It is frequently recommended because it provides sufficient space across the number of spindles required for handling a peak I/O load with today's larger disk drives. There are two physical I/O operations for each write requested (one to each mirrored disk).
  - RAID 5 (parity), which offers a higher usable capacity per RAID group then RAID 1/0. It can be effective for environments with very large mailboxes and/or lower IOPS requirements. There are four physical I/O operations for each write requested (two reads to calculate parity, one write for data and one write for parity).

BEST PRACTICE: Regardless of the constraints, the core requirements of providing enough drives to meet peak I/O demand remains, as does the strong recommendation to keep log and database LUNs for the same ESG on separate spindles.

Building-block design

It is recommended to create a clean storage layout diagram, using a building block style, assigning meaningful names to LUNs on an organized fashion. This will facilitate the understanding of the diagram, assist in identifying possible weaknesses, and eventually support the administration and maintenance of the implementation.

Design validation

Techniques such as peer reviews and comparisons against already proven configurations are recommended. Where possible, the configuration should be built and tested with performance tools such as Microsoft JetStress, LoadGen, and Performance Monitor, and EMC Navisphere® Analyzer.

Further detail on best practices for Exchange 2007 is available at *EMC CLARiiON Storage Solutions: Microsoft Exchange 2007 – Best Practices Planning*.

Specific technical documentation for Exchange solutions using EMC Symmetrix®, CLARiiON®, and Celerra® arrays is available at Microsoft Exchange Reviewed Program (ESRP).

Detail information for tools available to quantify an Exchange environment and sizing rules of thumb are provided in the *Symantec Enterprise Vault 7.0 - Sizing an Enterprise Vault Solution for Exchange Archiving – Technical White Paper*.

SQL Server

Enterprise Vault leverages SQL Server capabilities as its database engine. Due to the fact that Enterprise Vault performance heavily relies on the performance of the database server, SQL Server needs to be properly specified.

A holistic approach to application, host, and storage configuration planning is required to ensure optimal configuration for a given SQL Server database deployment.

Storage

Eliminating contention, a traditional best practice for database layouts for storage-related resources involves understanding how the database manages the data flow process and ensures that concurrent or near-concurrent storage resource requests are separated on to different physical spindles. Layout recommendations in this regard include:

- Ensuring that partitions, created on LUNs presented from storage arrays, are aligned to 64 KB boundaries during the file system partition creation phase. This partition alignment and the resulting volume alignment should not be confused with the Windows Allocation Unit size specified when formatting the volume. These are two different processes.

- Having transaction logs on separate hypers and spindles, to minimize contention for the logs as new writes come in from the database and any old transaction log information is streamed out during incremental transaction log backups. It also isolates the sequential write and random read activity for these members from other volumes with differing access characteristics.

- Isolating TEMPDB data and log files from other user databases and optionally allocating multiple data files (usually one per processor) for TEMPDB data by hosting them on separate LUNs.

- Implementing only a single file on any given LUN. In general, this provides for the best performance configuration and may not always be possible. Windows Server and HBA configurations create device I/O queues based on LUNs. Ensuring that queue structures are scaled out, and that workloads are optimized for a LUN, will in turn result in performance scaling.

- Utilizing multiple files within a filegroup to distribute I/O load. SQL Server will allow for certain parallel operations when tables have been created on multiple files. Full table scans would represent one such parallel operation.

Note: Although these recommendations may be considered to be best practices, in certain circumstances they may not be possible to implement. In such cases, it is possible to collocate these files in shared locations. However, this will require constant monitoring and management to ensure that the overall performance of the environment is not suffering as a result of these competing workloads.

Other considerations include the organization of the database data files to facilitate recovery when array replication technologies such as TimeFinder and SRDF are used. Co-locating data and log files or other files on LUNs (viewed as physical disks by Windows servers) will affect functionality. Specifically, an attempt to restore a database located on a LUN may result in the inadvertent and erroneous restoration of other unrelated data.

BEST PRACTICE: It is recommended in a SQL Server environment that separate LUNs only contain volumes that contain database files from the same database, or databases that will be backed up and restored together.

Applications availability and performance requirements determine the appropriate level of RAID to be configured. Although EMC recommends RAID 1 to be the primary choice in RAID configuration for reasons of reliability and availability, SQL Server databases can be deployed on RAID 5-protected disks for many database environments.

BEST PRACTICE: For Enterprise Vault environments, it is recommended to have:

- Transaction log files on RAID 1 or RAID 1/0 on 15,000 rpm (15k) drives, as they require support for high I/O demands.
- Data files on RAID 5 on 10,000 rpm (10k) drives, as read performance is as critical as rapid writes.

Further detail on best practices for SQL Server is available in the *Microsoft SQL Server on EMC Symmetrix Storage Systems TechBook*.

Host

There are many factors to consider when determining the number of Enterprise Vault servers that one SQL server can support. Table 1 shows the recommended number of Enterprise Vault servers that one SQL server supports.

Table 1 Recommended number of SQL Server servers

Enterprise Vault servers per SQL server	Recommended in this case
4	There is no maintenance plan, and the SQL server has a similar specification to the Enterprise Vault servers.
8	The databases are regularly maintained, or the SQL server has 8 GB of memory.

Note: Although possible, it is not recommended to host SQL Server on the same server where Enterprise Vault is installed.

SQL databases require periodic maintenance and as part of the maintenance plan, it is advised that the following actions should be taken on a weekly basis:

- Rebuild indexes
- Update statistics
- Shrink databases

Application

The databases required by Enterprise Vault are:

- Enterprise Vault Directory database - This database holds the configuration information for an Enterprise Vault site; usually there is only one Directory database.
- Vault Store database - There is a one-to-one relationship between a Vault Store and the database holding specific configuration information on the store and the metadata associated to each item stored in its archives.
- Vault Store Group fingerprint database - Enterprise Vault 8.0 introduced this database to support the new storage model. The new database has tables with entries for every single instanced part, holding fingerprints of the shareable parts of archived items.

Note: Although this database is not used when EMC Centera is the archiving storage device, it will be created and will be used in the event that other storage types are added later to the environment.

- Monitoring database - This database holds monitoring information for the Enterprise Vault Site.
- FSA Reporting database (optional) - When FSA Reporting is configured, this database holds the data that the Vault File Collector Service gathers.
- Audit database (optional) - When audit is enabled, this database holds all auditing events for the selected Enterprise Vault servers.

Additional information is provided in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

Database capacity requirements

Most of the required databases have predictable growth over time; however, Vault store and Vault store group fingerprint databases' growth depends on many factors unique to the organization as they will reflect the Exchange environment being archived.

For every item archived (managed) by Enterprise Vault, several attributes – usually referred to as metadata – are added to a vault store database; this is quite predictable once an assessment is performed. Temporary space is used to hold information on items that have not been backed up or indexed. Certain records are held longer for users who have been enabled for Vault Cache. Permanent space is also used to hold data in the Directory database. It is suggested that an extra 5 GB is allowed for this, or 10 GB where millions of items are archived between backups or Vault Cache is enabled for thousands of users. The extra space is added once only.

To calculate the space required for Exchange mailbox and journal archiving:

1. Take the number of items to be archived.
2. Multiply by 1,000 bytes (size of the metadata stored per item).
3. Add 5 GB.

Note: Calculations for the storage requirements of the Vault store group databases is more complex and is not part of the scope of this document as they are not relevant for Centera environments.

Table 2 lists the initial database file sizes for both data and transaction logs, as well as the recommended allocated storage per database to allow for temporary growth on a two-year projection.

Table 2 Recommended database storage allocation

Database	Recommended storage allocation	Initial database sizes (Data / Log)
Enterprise Vault Directory database	5 GB	10 MB / 25 MB
Vault Store databases	5 GB + (1,000 bytes * number of archived items)	100 MB / 80 MB
Vault Store Group fingerprint databases	Not required for EMC Centera	164 MB / 80 MB

Monitoring database	5 GB	100 MB / 80 MB
FSA Reporting database	5 GB	100 MB / 80 MB
Audit database	5 GB	100 MB / 80 MB

Further detail on SQL Server sizing is available in the *Symantec Enterprise Vault 7.0 - Sizing an Enterprise Vault Solution for Exchange Archiving – Technical white paper*.

Enterprise Vault

User profile

This information will help with understanding the profile for a set of users, over time, and what their archiving storage requirements are.

Mailboxes

- Total number of mailboxes to be archived (regular e-mail and journal) and size limits
- Mailbox distribution across multiple Exchange servers/locations
- Amount of messages to be archived (from backlog and each day) and size
- Anticipated growth over the next few years
- Retention policies

Archiving timeframes

- Expected time to clear the backlog
- Daily window timeframe for e-mail archiving
- Exchange journal schedules

Midterm considerations

- Mailbox size growth expectations
- Increased number of mailboxes over time
- How to deal with excess resources after the backlog is archived (if initially required)

Enterprise Vault Consulting and Systems Engineering teams use assessment tools, to be executed on customer premises, to determine the number and size of messages within a certain age band and consequently improve the accuracy of the calculations provided in the *Symantec Enterprise Vault 7.0 - Sizing an Enterprise Vault Solution for Exchange Archiving – Technical white paper*. Of particular interest in this paper are the typical values used in sizing estimates, which have been the guidelines used for all joint performance benchmarks.

A summary of this user profile is:

- Average message size is 75 KB.

- Typical user sends and receives are about 70 messages per workday.
- Typical age-based archiving policy is about 60 days.
- About 20 messages will remain in the users' mailbox and will be archived.
- About 30 percent of messages reaching archiving age are single instanced.
- Messages, including attachments, will compress between 30 percent and 50 percent.
- Average available archiving window is 6 hours, on a five-nights-per-week scenario.
- Prior to archiving, average backlog size is about 70 percent of a typical Exchange mailbox store.

In general, a dual-processor Enterprise Vault server can support mailbox archiving from up to 7,500 mailboxes (see Table 3).

Table 3 Supported mailboxes for mailbox archiving based on EV server configuration

Number of Enterprise Vault server - processors	Number for managed mailboxes
1	4,500
2	7,500
4	12,000

Index

Enterprise Vault uses AltaVista's Search Engine to fulfill end-user search requests, having one index per user or journal archive. As these AltaVista indexes are binary files with very high I/O demands, it is required that the computer hosting the Enterprise Vault Indexing Service has access to adequate storage for such indexes.

Indexes may be placed on local storage, SAN or NAS. However, if fast indexing is required or searches across a large number of archives is to be performed, it is recommended to use SAN devices instead; the same recommendation applies for fast concurrent searches generated by the use Enterprise Vault Discovery Accelerator or Compliance Accelerator in Enterprise Vault environments.

- Due to the high I/O demands on Enterprise Vault indexes, they should be treated the same way as SQL Server transaction log files, and RAID 1 or RAID 1/0 partitions on 15,000 rpm (15k) drives should be allocated.

Note: It is important to exclude the index locations from virus checking applications as anti-virus software can potentially change data.

Note the following about indexes:

- On NAS devices, opportunistic locking must be turned off.
- Indexes become fragmented, whatever the type of device and this slows down both searching and indexing. It is recommended to regularly defragment indexes, ideally while the Indexing Service is stopped, so that defragmentation does not conflict with updates. This is particularly important when the Accelerator products are being used. Additional information on how to tune the Accelerator products is available in the *Symantec Enterprise Vault 8.0 Performance Guide*.

Index storage requirements are calculated as a percentage of the size of the original data being archived, depending on the Indexing type needed, as listed in Table 4.

Table 4 Index storage requirements

Indexing type	Estimated storage requirements
Brief	3% of original data
Medium	8% of original data
Full	12% of original data (recommended)

Note: The percentages for Medium and Full will be less if there is little indexable content. This is often the case where large attachments such as MP3 or JPEG files are part of the environment.

Collections

Enterprise Vault provides two mechanisms for storing items in EMC Centera — with collections and without collections. A collection is up to 100 items or 10 MB of data.

When collections are used, items are first archived to a local staging area until the Enterprise Vault Storage Service process collects files in this area and stores them on EMC Centera. When Centera collections are enabled the Storage Service processes the item in two steps:

1. Item is archived to a temporary saveset file to the staging area (performed by the StorageArchive process)

2. Multiple temporary saveset files are collected and stored in a Collection Clip (performed by the StorageFileWatch process); when complete the temporary saveset files are deleted.

Note: If the primary EMC Centera system is offline when the Storage Service starts, the Storage Archive process does not perform step 1. If, however, the primary EMC Centera system goes offline after the Storage Service starts, the Storage Archive process performs step 1.

A particular Enterprise Vault server can manage multiple Vault Stores; a Vault Store may have multiple collections, but only one opened at any given time.

BEST PRACTICE: For performance purposes, enabling EMC Centera collections is the recommended choice.

Note: EMC Centera collections should not be confused with NTFS collections used when archiving to NT file systems.

BEST PRACTICE: Due to the high I/O demands required for collections, RAID 1/0 is the recommended RAID type configuration.

BEST PRACTICE: The recommended size for staging collections is 5 GB for normal operations; however, for contingency / disaster recovery purposes, it is recommended to allocate 50 GB per Enterprise Vault server.

Other Enterprise Vault storage requirements

Table 5 presents two items to be hosted on each Enterprise Vault Server local disk. Although the amount of local storage is small, I/O performance is critical to avoid any impact on the application.

Table 5 Additional storage requirements per Enterprise Vault Server

Item	Purpose	Size
Temporary files	Used by Enterprise Vault archiving processes, in particular the Storage Service during archiving and conversion	100 MB
MSMQ files	Used during Exchange archiving and journaling	10 MB

Note: It is recommended to place these items on disks separate from the system disk, having RAID 1/0 as the RAID type configuration.

Note: It is recommended to reassign the TEMP system variable to a drive other than the C: drive.

Note: During installation, Enterprise Vault requires 70 MB of disk space to install all the Enterprise Vault components.

EMC Centera

Properly sizing the EMC Centera for optimal capacity and performance involves several factors. Some major factors would include:

- Estimated thread consumption
- Average object size
- Bandwidth
- Daily ingest rates
- Protection scheme
- Retention policies
- Capacity growth
- Node object count threshold
- Read write requirements.

During the Enterprise Vault archiving process, an item is first compressed and then metadata is added to it. As a general rule, the item is compressed to half its original size and the metadata compresses to approximately 5 KB.

Note: Compression ratios may vary considerably depending on the type of file. Office 2003 documents tend to compress well. Other document types, such as ZIP files or JPG files, are already compressed and cannot be compressed further.

Note: When archiving to EMC Centera, Vault Store Partitions, which are actually the only things stored on the EMC Centera, still use the pre-EV8.0 OSiS format, as Enterprise Vault relies solely on EMC Centera single-instancing capabilities.

As a rough rule of thumb, it is recommended to assume that the storage required on the EMC Centera (not including the protection mechanism – CPM recommended) will be approximately 50 percent of the original unarchived size. See the *Symantec Enterprise Vault 8.0 Performance Guide* for more details on the EMC Centera sharing (single instancing) model.

BEST PRACTICE: In addition to data growth, protection schema, potential single instancing, retention policies, and other legal requirements greatly impact the size of the archive. These factors should be carefully considered.

Note: EMC Centera single instancing is achieved across all Enterprise Vault's store partitions, regardless of the Vault Store Group they belong to. This background task, transparent to the application, provides further storage savings, and minimizes the network planning requirements.

Network

Although the network infrastructure seldom is the limiting factor on performance, in order to support the considerable volume of network traffic potentially generated by Enterprise Vault, it is recommended to provide an environment in which the connections support the expected response time of at least 100 Mb/s switched Ethernet LAN. Usually, a 100-Base-T is sufficient, but unique requirements and constraints require more precise calculations on network usage for the various archiving components.

Note: For guidelines on the network traffic expected between the various components under different conditions, it is recommended to contact the Symantec supplier responsible for the deployment. The *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* has more information

EMC Centera network segmentation allows for the filtering of network traffic by use case (application data access, Centera Replication, Centera management). This provides not only extra flexibility during the implementation of the solution but also room for network traffic optimization and fine tuning.

BEST PRACTICE: Although optional, for environments with high network traffic, it is recommended to implement EMC Centera network segmentation. Additional information is provided in the *EMC Centera Network Segmentation* white paper.

Infrastructure / business continuity considerations

During the planning phase, it is recommended to identify the critical activities or functions that are part of the e-mail archiving solution and determine their recovery point objective (RPO) and recovery time objective (RTO) in order to ensure that business or technical requirements are met in case of a disaster recovery event.

Enterprise Vault and EMC Centera provide several mechanisms that assist enterprises in minimizing the risks associated with these events and properly recover when they do occur.

Enterprise Vault backup

Given Enterprise Vault's flexibility and scalability, its components can be distributed across multiple computers, requiring an effective backup strategy to prevent data loss, and to provide a means for recovery in the event of a system failure. In particular, Enterprise Vault Services and tasks may be remote to the resources Enterprise Vault depends on.

To facilitate backup activities and minimize impact on end users, Enterprise Vault provides a backup mode setting for Vault Stores and index locations. When set up, Enterprise Vault backup mode places the Vault Store/index location in read-only mode while the backup is taken, ensuring that users are serviced for searches and retrievals from the archive while the data is in a steady state.

The *Symantec Enterprise Vault 8.0 - Administrators Guide* has additional information on how to set/clear the Vault Store and index locations' backup mode, including examples for typical use cases.

Data to be backed up

A complete system and file backup, including the registry – because all Enterprise Vault Services store information in the registry – is required. It is recommended to take this system and file backup at the same time the following Enterprise Vault system databases are being backed up:

- EnterpriseVaultDirectory
- EnterpriseVaultMonitoring
- EnterpriseVaultAudit
- EnterpriseVaultFSAReporting

Finally, each vault store database must be backed up, as well as all the index locations for that particular Vault Store.

Note: When archiving to EMC Centera, as opposed to any other storage device, it is not required to back up the vault store partitions.

The *Symantec Enterprise Vault 8.0 - Administrators Guide* provides additional information on the following topics:

- Enterprise Vault backup requirements, and procedures to determine the locations of the several components
- Enterprise Vault failover configuration
- Enterprise Vault recovery procedures

Note: It is recommended to use a building block approach to setting failover, as it is the practice of backing up the complete Enterprise Vault environment, as listed here.

The *Enterprise Vault 8.0 and later - Backing up Enterprise Vault* document contains detailed information about backing up Enterprise Vault, including details on how to create a view and store procedures to automatically back up all Enterprise Vault databases, using backup scripts.

Additional information on performing backups of Enterprise Vault systems that are clustered, using Veritas Cluster Server (VCS) or Microsoft server clusters, is available in the *Backing up Enterprise Vault in a clustered environment technical note*.

Enterprise Vault clustering

As with any other application critical to the business, clustering of an Enterprise Vault environment provides a means for a high-availability solution. Another alternative is to implement Enterprise Vault building blocks.

For clustering purposes, Enterprise Vault can be integrated with Veritas Cluster Server (VCS) or in Microsoft server clusters.

Note: Active/Active clustering configurations are not permitted by Enterprise Vault.

On a VCS environment, active/passive and N+1 configurations are supported. The VCS GenericService agent leverages the information contained in the Enterprise Vault Directory database to manage the following resources:

- Admin Service
- Directory Service
- Indexing Service
- Shopping Service
- Storage Service
- Task Controller Service

Additional information on VCS configuration is available in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

On a Microsoft server cluster environment, the following operation modes are supported:

- An active/passive failover pair - A primary node with a dedicated failover node.
- N+1 (hot standby server) - Two or more primary nodes share a single failover node. Only one node failure can be accommodated at any one time.
- N+M - An extension of the hot standby concept with N primary nodes and M failover nodes. Only M node failures can be accommodated at one time.
- N+M “any-to-any” - Identical to N+M, except that there is no need to fail back to the original node after a failover. When the original node becomes available again, it can operate as a failover node.

The *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* has additional information on how to configure Enterprise Vault with Microsoft server clusters.

For failover capabilities, configuring Enterprise Vault following a building block approach is a feasible alternative as it provides a means for having a solution that is both reliable and scalable. Each building block comprises a full set of Enterprise Vault Services that deliver archiving, indexing, storage, search, and retrieval functions.

Note: Building blocks in a clustered configuration are not currently supported by Enterprise Vault.

The *Symantec Enterprise Vault 8.0 - Introduction and Planning Guide* has additional information on possible configurations.

Centera Replication

EMC Centera Replication provides a disaster recovery mechanism for content written to EMC Centera clusters. EMC Centera Replication can be used to create multiple protected copies of content written to a primary Centera cluster by copying the content to a replica Centera cluster. Replication runs as an asynchronous background task and can be configured in a number of topologies, to best fit customer requirements.

EMC Centera Replication provides functionality that allows the recovery of content that is either missing or simply unavailable. Applications will fail over automatically to retrieve missing or unavailable content on the primary cluster from the replica cluster.

The main characteristics of Centera Replication are:

- The process automatically copies new content to another EMC Centera cluster. No backups of archived data are necessary.
- The Centera Replication mechanism ensures that new content from a local application is automatically and transparently transferred across a WAN to a designated EMC Centera cluster, presumably in another location.
- Provides a disaster recovery mechanism for content written to EMC Centera clusters.
- Can be used to create multiple protected copies of content written to a primary EMC Centera cluster by copying the content to a replica Centera cluster.
- Provides functionality that allows the recovery of content that is either missing or simply unavailable. Applications will fail over automatically to retrieve missing or unavailable content on the primary cluster from the replica cluster.
- Replication runs as an asynchronous background task and can be configured in a number of topologies. These combinations provide the replication topology that best fits customer requirements:
  - Unidirectional
  - Bidirectional
  - Chain
  - Star

Note: Enterprise Vault indexes and SQL databases are not held on EMC Centera and still require backups.

In a typical replication setup the EMC Centera clusters are geographically separate to ensure disaster recovery or to distribute the content for access from another location. For example, a company may replicate to a second EMC Centera cluster to enable recovery from the loss of the primary Centera or to avoid multiple requests for the same content across a WAN connection.

Unidirectional replication

The most basic form of replication is unidirectional replication between two EMC Centera clusters. The application writes data to cluster A and that data is automatically copied to cluster B (active passive configuration).

In case of a disaster or when the primary EMC Centera cluster becomes unavailable, the application may failover to the replica cluster B. Automatic read failover is a feature of the EMC Centera SDK and is enabled by default, but other failover options may be configured (see Figure 2).

Note: Data written to cluster B during a disaster needs to be restored to cluster A, once it is back online. If cluster A was lost during the disaster, all lost data needs to be restored from cluster B, using the EMC Centera Restore feature. To avoid having to restore the content back, it is recommended that bidirectional replication be set up to capture all new content.

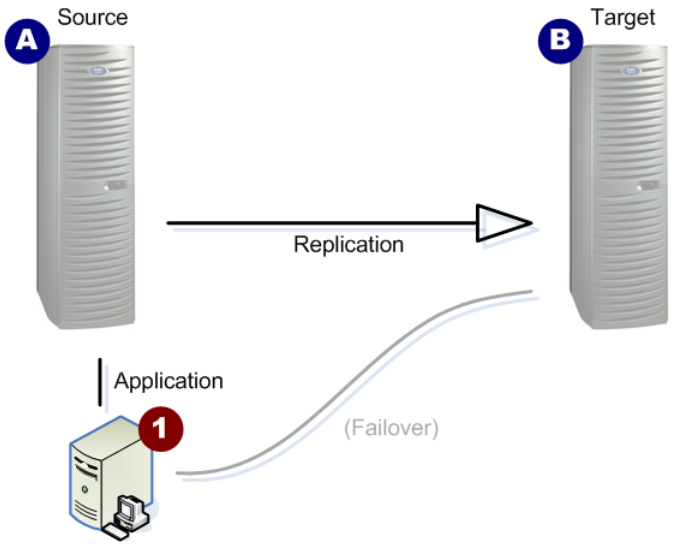

Figure 2 Replication failover

Bidirectional replication

The bidirectional replication topology allows applications to write to either of the EMC Centera clusters and have data copied to the other Centera cluster. Data written to cluster A will be replicated to cluster B and vice-versa.

This allows for establishing a complete DR solution whereby the application servers are also redundant. When the complete site A is lost, site B can take over immediately. This is called hot standby and is shown in Figure 3. Data written to cluster B by application 2 or application 1, if it was cut over to site B during a disaster, will be held in a replication queue and replicated to site A as soon as this site comes back online. If site A was lost during the disaster, all lost data must be restored from cluster B, using the EMC Centera Restore capability.

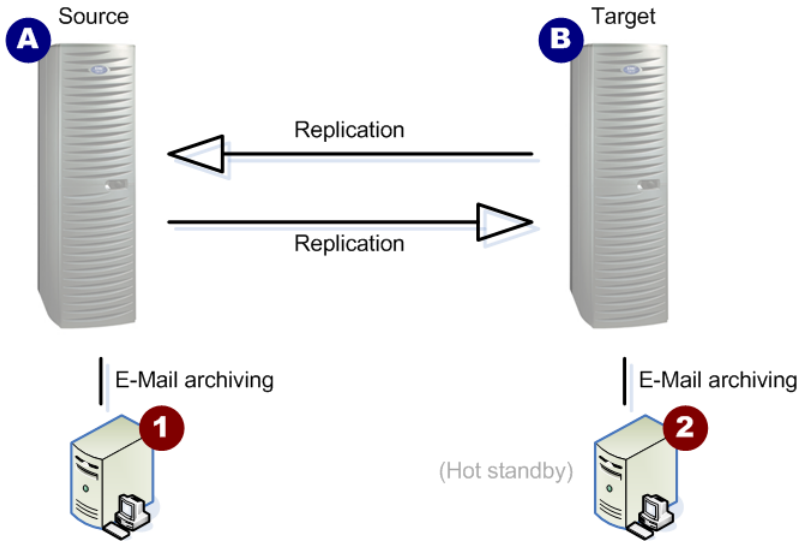

Figure 3 Bidirectional replication with hot standby

Rather than using site 2 as a hot standby, both sites can be used for production. This is useful when both sites have users who work on the local application server. All data will be kept available at both sites using the bidirectional application.

Note: Although from an EMC Centera perspective all data can be shared between sites, each local application may not be aware of data created by the other application. Enterprise Vault databases may need to be synchronized.

Additional information on Centera Replication, including other topologies available, and Restore operations, is available in the *EMC Centera Replication – A Detailed Review* white paper.

BEST PRACTICE: Content on the primary (source) cluster that was not yet replicated to the secondary (target) cluster may not be recoverable. Ensuring that the Enterprise Vault Remove safety copy feature is enabled will seamlessly allow the re-archival of the data, preventing any potential data loss.

Centera Restore/tape out

To copy data from one cluster to another, EMC Centera offers restore in addition to replication. Replication is an ongoing process that starts after it has been set up and continues unless it is paused by the system administrator or by the system. Restore is a single operation that copies

data from a source cluster to a target cluster and is only performed as needed by the system administrator. A restore operation can be performed from any source cluster to any target cluster. There is no need for a replication setup between the two clusters.

The main characteristics of the Restore capabilities are the following:

- Restore is a functionality of EMC Centera which, when enabled, copies data (clips, BLOBs and mutable metadata) from a source cluster to a target cluster. The set of data to copy can be narrowed down according to the use case at hand, full or partial.
- Restore is a process complementary to replication, whereby an EMC Centera cluster copies some set of its content or the entire Centera to another Centera cluster.
- Replication is used on an ongoing basis to keep two EMC Centera clusters synchronized with new content, whereas restore is used only as needed to populate one EMC Centera cluster with the content of another Centera.
- Restore is capable of restoring content at the cluster or pool level.

Note: When business requirements determine that a copy of EMC Centera data on tape or Virtual Tape Library (VTL) is needed, alternatives to perform such activities are available with or without the use of NDMP technologies. Additional information on available products to support these activities is available through Symantec or EMC.

Chapter 2 Implementation

This chapter presents these topics:

Installation requirements and configuration 38
Installation process 54
Configuring Enterprise Vault 57
Installation validation 80

Installation requirements and configuration

Enterprise Vault

Before installing or upgrading Enterprise Vault, it is strongly recommended to run the Enterprise Vault Deployment Scanner to check prerequisite software and settings to review the configuration of a computer and to report on any issues that may stop Enterprise Vault from running on it. The *Symantec Enterprise Vault 8.0 - Deployment Scanner Guide* provides a list describing the tests that Enterprise Vault Deployment Scanner performs, as well as directions on installing the tool.

Enterprise Vault Deployment Scanner is a separate wizard that is supplied on the Enterprise Vault media. When the tool runs, it creates a Reports folder in the folder in which it is run, and places a report file in the Reports folder.

Deployment Scanner and accompanying documentation may be found in the Enterprise Vault 8.0\Deployment Scanner folder on the Enterprise Vault media.

Note: Windows Installer 3.1 must be installed on your Enterprise Vault servers in order to install Enterprise Vault Deployment Scanner and the Enterprise Vault server components.

Software requirements

This section describes the operating system and software requirements for the core Enterprise Vault Services.

Note: There may be additional requirements for the different types of archiving. The requirements for Exchange are described in the Exchange Archiving with Enterprise Vault section of this document.

Operating system

Enterprise Vault server components can be installed on the following operating systems:

- Windows Server 2003 with Service Pack 2 or later
- Windows Server 2008 with Service Pack 1 or later

If Windows Server 2008 with Service Pack 1 is installed, the following mandatory hotfix for IIS must also be installed, as indicated by Microsoft at http://support.microsoft.com/kb/949516.

For additional details of supported versions, see the Symantec Enterprise Vault (tm) 6.0, 7.0, 2007 and 8.0 Compatibility List.

Install Windows with the following options and components; additional information is listed in the pages that follow, as well as in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

- NTFS file system
- Microsoft Message Queuing (MSMQ) Services
- .NET Framework 2.0
- Internet Information Services (IIS) 6.0 or later

- Internet Explorer 6.0 or later
- MSXML

MSMQ

Enterprise Vault tasks use MSMQ to communicate with the Storage Service. To install Enterprise Vault Services on more than one computer in the network, MSMQ must be configured on each computer.

The steps for installing MSMQ on Windows 2003 and 2008 are different. Note that Active Directory Integration should not be enabled when installing MSMQ.

Additional info on how to install MSMQ on different platforms is available in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* according to the operating system of use.

.NET Framework

It is required to install Microsoft .NET Framework 2.0 on Enterprise Vault servers.

If necessary, it can be downloaded using the link in the **Links to related software** folder on the Enterprise Vault media or directly from Microsoft .NET Framework Version 2.0 Redistributable Package (x86).

Note: NET Framework 3.0 SP1 or later is required for Compliance Accelerator (version 8.0 and later) and Discovery Accelerator (version 8.0 and later). See the *Symantec Enterprise Vault (tm) 6.0, 7.0, 2007 and 8.0 Compatibility List* for more details.

IIS

It is required to install IIS 6.0 or later on each Enterprise Vault server.

In IIS, it is possible to configure the level of isolation for particular Web applications. For shopping baskets in the Enterprise Vault Web Access application to be created correctly, the application needs to run under the predefined Local System account.

The configuration wizard will automatically set the correct isolation and account settings; consequently, there is no need to configure this.

If IIS 6.0 or later is already installed, the configuration wizard will create a new Application Pool, EnterpriseVaultAppPool, for the Web Access application and assign the Local System account to that pool.

Note: Active Server Pages and ASP.NET must be enabled. On Windows Server 2003, it is required to enable Active Server Pages and ASP.NET manually. The *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* has more information.

In Windows Server 2008, if the Add Roles wizard is used to install IIS, the wizard will provide the default installation, which has a minimum set of role services. Minimum IIS-related roles services required for Enterprise Vault are listed in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

Internet Explorer

Internet Explorer 6.0 or later is recommended.

MSXML

All Enterprise Vault server computers require MSXML. This is installed automatically with Internet Explorer 6.0 and later.

If an earlier version of Internet Explorer is being used, MSXML may be required to be installed. This is available from a link in the **Links to related software** folder on the Enterprise Vault media.

Roles-based administration

Roles-based administration uses Microsoft Windows Authorization Manager. Creating and managing roles using the Administration Console requires the Authorization Manager MMC snap-in, which is only available on the following:

- Windows Server 2003
- Windows Server 2008
- Windows XP Professional, or Windows Vista with the Administration Tools Pack for Windows Server 2003 and Windows Server 2008

The Administration Tools Pack can be downloaded from the following location:

http://www.microsoft.com/downloads/details.aspx?FamilyID=c16ae515-c8f4-47ef-a1e4-a8dcbacff8e3&DisplayLang=en

SQL Server

Enterprise Vault supports SQL Server 2000, 2005, and 2008 SP1 as database engines. Additionally, both **Windows Authentication** mode and **Mixed Mode Authentication** are supported.

Note: The SQL installation must be case-insensitive, as case-sensitive SQL installations are not supported.

When installing both Enterprise Vault and SQL Server on the same Windows Server 2003 computer, at least SQL Server 2000 with Service Pack 4 is required. Windows Server 2008 requires SQL Server 2005 SP2 or later.

Microsoft Data Access Components (MDAC)

To enable access to the SQL databases, MDAC 2.6 or later must be installed on Enterprise Vault servers. A suitable version is installed automatically with both Windows Server 2003 and Windows Server 2008 (which changes the name of MDAC to Windows Data Access Component or Windows DAC).

If necessary, the software can be installed using the link supplied in the **Links to related software** folder on the Enterprise Vault media. The *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* has additional information.

Hardware / Storage requirements

The most critical factor in the performance of Enterprise Vault is the specification of the system that is used—the servers, the disk systems, the memory, and the network.

Application server

In general, the more powerful the processor, the higher the ingest and retrieval rates. The other components of the system—the disks, memory, network, and archiving source—need to match this power. The following represents general considerations on server selection and configuration:

- Enterprise Vault must be installed on a computer member of a domain.
- Enterprise Vault makes good use of multi-core processors, and quad-core processors are recommended.
- 4 GB of memory is recommended. The operating system boot flag /3GB must not be used as this does not provide any benefit and can result in running out of system page table entries. For a fuller explanation, see "32-bit and 64-bit platforms" in the *Symantec Enterprise Vault 8.0 Performance Guide*.
- Turning hyperthreading on or off makes no noticeable difference to the overall performance of Enterprise Vault, so it is recommended that the manufacturer's default setting is not changed.
- There is no performance difference when running Enterprise Vault on 64-bit Windows (WOW 64) when compared with 32-bit Windows. For a fuller explanation, see "32-bit and 64-bit platforms" in the *Symantec Enterprise Vault 8.0 Performance Guide*.
- There is no performance difference when running Enterprise Vault on a Windows 2008 operating system when compared with Windows 2003 or when archiving from an Exchange Server on a Windows 2008 system.

Table 6 shows the recommended minimum specifications for a production Enterprise Vault system supporting 7,500 user mailboxes.

Table 6 Minimum specification for an Enterprise Vault server

Item	Recommended minimum
Number of CPUs	2
Power of CPUs	2.8 GHz
Memory	4 GB

It is possible to run Enterprise Vault on a computer with less memory, but this is not recommended for a production system, as it does not allow for any growth in archiving requirements. The extra memory is particularly important if users will be performing large, simultaneous archive searches.

Enterprise Vault can be run on a multi-processor system with four or eight CPUs, but in order to take advantage of the extra CPU power, the disk system used must be able to cope with the increased throughput.

In a small to medium Enterprise Vault environment, the core Enterprise Vault Services will typically all be installed on the same computer. In larger installations, services such as the Storage and Indexing Services can be installed on a separate computer.

The *Symantec Enterprise Vault 8.0 - Introduction and Planning Guide* has more information on distributing Enterprise Vault Services.

Additional processing capacity for initial archiving

If a large backlog of data that needs to be archived quickly, when Enterprise Vault is installed, additional Enterprise Vault servers may be configured for the initial archiving run. When archiving reaches a steady state, the additional Enterprise Vault servers can be redeployed for other purposes.

Database server

The recommended specification for best performance is a standard SQL server with four physical CPUs and a minimum of 4 GB of RAM; for medium or large environments, 8 GB of RAM is recommended. It is also recommended to fine-tune the performance of the SQL Server using standard methods, such as the ones described in the *Symantec Enterprise Vault 8.0 Performance Guide*.

- Using an x64-based 64-bit platform provides more efficient memory utilization and brings performance benefits. The appropriate edition of Windows Server 2003 and SQL Server 2005 must be installed to support the capacity of memory installed, but no other tuning options need to be set.

- If a 32-bit database server will be used then the server should be carefully tuned to make best use of available memory. These tuning options depend upon using the appropriate edition of Windows and SQL Server for the installed capacity of memory. If the database server has more than 4 GB of physical RAM:

  - Enable the operating system Physical Address Extensions boot flag (/PAE).

  - Enable Address Windowing Extensions (AWE) memory in SQL Server using the following script:

```
sp_configure 'show advanced options', 1
RECONFIGURE
GO
sp_configure 'awe enabled', 1
RECONFIGURE
GO
```

- Hyperthreading may not be beneficial to SQL Server environments and should be carefully tested before enabling.

Enterprise Vault requires a number of SQL databases:

- The Enterprise Vault Directory database holds the configuration information for an Enterprise Vault Site.

- Each Vault Store has a Vault Store database, which holds configuration information for the Vault Store and details of the items stored in its archives.

- Each Vault Store Group has a fingerprint database, which holds the fingerprints and other information related to the single instance storage parts that are created for Enterprise Vault single instance storage.

- The Monitoring database holds monitoring information for the Enterprise Vault Site.
- If FSA Reporting is configured, Enterprise Vault creates an FSA Reporting database to hold the FSA Reporting data.

The SQL Server that manages these databases will typically reside on a different computer from the Enterprise Vault server.

In general, the specification of the SQL Server computer should match that of the Enterprise Vault server. The performance of the SQL Server will also benefit from extra memory; a minimum of 4 GB is recommended. The amount of memory that the SQL Server can use depends on the Windows and SQL Server versions. Table 7 shows the recommended minimum specifications for a production SQL server.

Table 7 Minimum specification for SQL Server

Item	Recommended minimum
Number of CPUs	2
Power of CPUs	2.8 GHz
Memory	4 GB

Note: There is no need to have a separate SQL server for every Enterprise Vault server. As a general rule, one SQL server can manage four to eight Enterprise Vault servers. For additional information, see the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

Indexes

Note the following points:

- It is required that the default index storage location is on an accessible device and that the Vault Service account can write to it.
- With Exchange Server archiving, Enterprise Vault adds information about the index storage location to the Directory database once the mailboxes are enabled. It is recommended to perform any changes to the index storage location, or add further locations before any mailboxes are enabled via the Administration Console. Changing the index storage location for mailboxes after they have been enabled cannot be done easily.
- As anti-virus software can potentially change data, it is important to exclude the cache and index locations in the virus checking application. The *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* has more information.

Exchange archiving with Enterprise Vault

This section covers additional requirements and considerations for Microsoft Exchange archiving using Enterprise Vault.

The following target Exchange servers are supported for archiving items from mailboxes and public folders:

- Exchange 2000 with Service Pack 3

- Exchange Server 2003
- Exchange Server 2007 with the Mailbox Role installed

Outlook 2003 is required on any Enterprise Vault server performing Exchange Server archiving tasks.

In addition to the previous requirements, the following pre-installation tasks are required prior to installing and configuring Exchange Server archiving components. The *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* has detailed information.

The Enterprise Vault system mailbox

A system mailbox is required on each Exchange server that Enterprise Vault is to archive; this Enterprise Vault system mailbox will be used by the Exchange mailbox, Exchange journaling, and Exchange public folder tasks when connecting to the Exchange server.

This mailbox name is required by the Administration Console whenever an Exchange Server archiving task is created. The following restrictions apply to this mailbox:

- The Enterprise Vault tasks require exclusive use of this mailbox, so the mailbox must not be used for any other purpose.
- The mailbox must not be hidden from address lists.
- The account must not be disabled

Note that after creating the Enterprise Vault system mailbox, it may take some time for the mailbox to be available. The amount of time depends on configuration options in Exchange System Manager. The mailbox must be available before an Exchange Server archiving task is added. Additional information on how to manually force a mailbox update on Exchange Server 2000 or 2003 is available in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

Additional Vault Service account permissions

The Vault Service account must be a member of the Active Directory domain. It is recommended to assign Exchange Server permissions explicitly, as detailed in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* rather than making this account a Domain Administrator.

Assigning permissions on Microsoft Exchange Server

Some of the permissions required are as follows:

- Microsoft Exchange 2007 with the **Mailbox Role** installed
  - On Active Directory, add the Vault Service account to each Exchange Server 2007 server object, granting **Full Control** to **This object and all child objects**
  - **Send As** permissions are required.
- Microsoft Exchange 2000 or 2003
  - On Microsoft Exchange System Manager, add the Vault Service account, ensuring the all **Allow** checkboxes are selected.

Create an Outlook profile on the Enterprise Vault server computer

When Outlook is installed on the Enterprise Vault server, a profile must be created and a connection to an Exchange Server mailbox must have been established before installing Enterprise Vault.

Configure Internet Explorer

Microsoft Office Outlook must be set as the default e-mail application for Internet services in Internet Explorer on the Enterprise Vault server.

Other Enterprise Vault components

- Stand-alone Vault Administration Console
- Operations Manager
- Discovery Accelerator
- Compliance Accelerator
- EMC Centera Checker
- EVR

EMC Centera

This section presents an overview of the features offered by EMC Centera applicable to an environment where Enterprise Vault is to be deployed, and the recommendations to configure EMC Centera to best perform in such setting.

The recommended EMC Centera configuration for Enterprise Vault is listed in Table 8.

Table 8 Recommended EMC Centera configuration

Feature	Recommendation
Content Protection	Mirrored (CPM) (default)
Storage Strategy	Performance Full (default)
Pool/Profile	Enterprise Vault requires a profile to have Read, Write, Delete and Exist capabilities
PEA File	Use of a Pool Entry Authorization (PEA) file is recommended
Replication	At least Unidirectional is recommended; other topology depending on business requirements
Compliance Mode	At least Governance Edition; Compliance Edition recommended based on business requirements

Note: Content Protection and Storage Strategy are values set at the time of EMC Centera installation; it is recommended to verify that the EMC Change Control Request Form accompanying the box has the appropriate values.

Content Protection

EMC Centera provides two protection schemes for storing data objects redundantly:

- Content Protection Parity (CPP)
- Content Protection Mirrored (CPM)

For additional information, see the *EMC Centera Programmer's Guide*.

Content Protection Parity (CPP)

CPP or parity is the process whereby each stored object is split into six fragments. Each fragment will be stored on a different node in the same cluster. CPP calculates a parity fragment from the stored fragments and stores that as the 7th data fragment on yet another node. This provides the ability to reconstruct the object in the event of data loss of any one of the seven fragments.

CPP is a more space-efficient way to store data at the cost of lower performance.

CPP segments a data object into six parts and stores each one on a different node in the same cluster. CPP calculates a parity fragment from the stored fragments and stores that as the 7th data segment on yet another node. This mechanism is a derivative of RAID 3 (mostly used in single-user systems with large record applications). If seven nodes are not available, data is stored with CPM only.

When CPP is enabled, not all files will be using CPP. CDFs are always written in CPM as well as small files. EMC recommends using CPP only for files greater than 250 KB.

The threshold that is set for CPP can impact the overall storage capacity expectations. If the majority of the stored files fall below this threshold then they will not be stored in CPP and the capacity benefits of CPP will be lost.

Content Protection Mirrored (CPM)

CPM or mirroring is the process whereby each stored object is copied to another node in an EMC Centera cluster. There is a dual power supply to each node, helping to ensure that at least one copy of the data will always be available in the event of a disk, node, or power failure.

CPM stores a complete copy of the data object on a mirror node. CPM enhances faster data regeneration and improves performance during normal operation. Object retrieval is improved because the access node with the access role can select the least loaded node with the storage role from which to retrieve a specific object.

CPM is the default protection scheme for storing C-Clips that are 250 KB or less.

Note: Additional information on EMC Centera content protection mechanisms, regeneration, and self-healing is available the EMC Centera online help.

Note: When CPP is enabled, a fallback can be configured for when there are not enough nodes available in any cube for writing the fragments on seven different nodes. The default fallback is to return an error to the application and disallow further ingest to the cluster. The system can also be configured to fall back to CPM in such a case.

Additional information for data protection and self-healing EMC Centera capabilities are available in the Centera online help.

BEST PRACTICE: Content Protection Mirroring is the default Centera protection scheme and is recommended when used with Enterprise Vault.

Storage strategy

EMC Centera provides two storage strategy mechanisms:

- Storage Strategy Capacity
- Storage Strategy Performance Full

Storage Strategy Capacity ensures single instancing – the archival of a single copy of an item, regardless of the number of references (instances) of the content in the original data source – and thus optimizes capacity.

Storage Strategy Performance Full improves the speed of write operations at the cost of single instance storage. When Storage Strategy Performance Full is enabled, identical content may be stored multiple times.

Note that setting the Storage Strategy to Performance Full does not improve performance substantially for large files (> 1 MB). When setting the Storage Strategy to Performance Full the default threshold (maximum file size) is 250 KB. Objects larger than this threshold will automatically revert to Storage Strategy Capacity in order to benefit from single instance storage. The threshold can be adjusted to optimize the balance between performance and capacity.

For CPM the threshold of 250 KB equals the file size. For CPP the threshold corresponds to a file size of 6*250 KB = 1.5 MB.

BEST PRACTICE: The Storage Strategy is set to Performance by default and is the recommended setting when used with Enterprise Vault, as it enhances Enterprise Vault Sharing capabilities for large objects, which usually represent the bulk of the archive.

Storage optimizations – Embedded BLOBs and containers

In addition to these storage strategies, EMC Centera provides other programmatic mechanisms to allow applications to further improve their integration and performance when archiving small items to EMC Centera:

- BLOB Embedding

  Very small data BLOBs (less than 100 KB) can be converted to a Base64 string and stored in the CDF, either as an embedded Data BLOB or in a Tag string attribute, as opposed to a BLOB object, hence non-impacting the object count and allowing the access to the data without incurring in the extra overhead of reading the data from EMC Centera.

- Containerization

  Another approach for managing small objects is to aggregate them into a single BLOB. This technique is called containerization. For example, many Call Detail Records (CDRs) can be taken from a telecom switch and stored as a single EMC Centera object. The offset and length for each component object can be stored as an attribute associated with the BLOB tag. Specific API calls (FPTag_BlobReadPartial) can then be used to retrieve the individual component objects from the container.

The following considerations are relevant to data containerization:

- In general, aggregated data does not benefit from single-instance storage. However, Enterprise Vault does archive attachments and other large objects individually (outside the container) in order to favor SiS. The retention period applies to the aggregated data, as all data archived in the container follows the same retention policies.

- Individual data objects cannot be deleted without rewriting the container object.

Note: If the requirement is to store many small files (< 50K), EMC recommends that the application embeds the file directly in the CDF or combines multiple small files in one C-Clip™. Refer to the *EMC Centera API Reference Guide* for more information on embedding data.

Enterprise Vault fully leverages both BLOB embedding and containerization capabilities. Small items are stored in the CLIP directly. Items that are smaller than 15 KB will be stored in the CLIP and since this is "relational" it tends to be faster. Enterprise Vault "knows'" how to store the item in the CLIP and retrieve it. Collections (or Centera containers) are done differently to NTFS. Taking e-mail as the example, only collected items will be sent to EMC Centera.

E-mails are split into three parts, per user, body, and attachments. If an item is over the 15 KB threshold it will be passed to the Enterprise Vault Collector Service – if collections are enabled in Enterprise Vault. The per-user and bodies are collected into larger collections, while the attachments (greater than 50 KB) are stored separately. For additional information see the Sharing section in this document. Using collections, Enterprise Vault not only improves its SiS efficiency but also improves the performance of EMC Centera, by sending fewer, larger files.

The body and per-user details are always going to fail SiS checking as they are by their nature unique. By sending them as part of the collection, it is ensured that Centera SiS testing is kept to a minimum, thus improving Centera throughput. This is another reason why setting the storage strategy to Performance Full (SSP) perfectly fits the model.

The details of the actual items that are stored in Centera (header, body and attachments) are then referenced in the original CLIP. Enterprise Vault will then read the CLIP content to see how to recover any piece of content based on the offset and size of the message, without having to retrieve the whole container. Along with all the information stored in the CLIP are Enterprise Vault generated checksums. These checksums help Enterprise Vault guarantee that the content committed to Centera is valid when recalled.

In general, these are the three primary factors in archiving with Enterprise Vault and EMC Centera:

- Saveset collection limit is 10 MB or 100 savesets.

- Anything larger than 50 MB is not compressed.

Note: As opposed to e-mail, NTFS or CIFS files being archived by Enterprise Vault FSA (File System Archiving) are handled differently. If the file is less than 50 MB it is handled like an e-mail message. If the file is larger than 50 MB it is stored in a different way. The file is not compressed, the contents are not indexed, and the saveset is always stored in an individual Saveset Clip (never as part of a Collection Clip).

BEST PRACTICE: Containerization is a storage technique strongly recommended by Centera which can be enabled or disabled (default) in Enterprise Vault. It is a best practice to enable Enterprise Vault Sharing (Centera containers) while setting up archiving in Enterprise Vault.

Virtual pools and profiles

Virtual pools (VP) allow storage administrators the ability to logically segregate application data stored on EMC Centera. In pre-CentraStar® 3.0 incarnations, CentraStar presented a single, communal pool that housed all data stored on a cluster. The value of the VP model is that it presents a highly abstracted method of operation that protects applications from details of the underlying storage mechanics. VPs represent a natural evolution of the pool model. The advantages of subdividing the communal pool are substantial and include:

- The ability to perform CentraStar operations on select subsets of data
- The ability to replicate on a pool basis
- Segregation of application data on the cluster
- The ability to set up hard pool capacity quotas
- Capacity reporting by pool
- The ability to restrict what IP can access a virtual pool

The exposed use of VPs is in data segregation and the ability to change the scope of the CentraStar Replication/Restore functions from the cluster at large to a set of discrete VPs. VPs allow you to change the scope of most cluster functions away from the global cluster to individual pools and they also make it transparent to the application accessing the EMC Centera storage. For additional information, see the *EMC Centera SDK Version 3.2 API Reference Guide*.

A Centera profile specifies what capabilities (or operations) can be performed on a connection established with that profile. The capabilities are Read, Write, Purge, Delete, Exist and Query.

A profile can have access to one or more pools, each pool/profile pair having a separate set of granted capabilities; consequently, it is possible for an application to establish a connection to EMC Centera for which there are Read, Write and Delete capabilities to one pool, but only Read privileges are given to another. This characteristic is particularly relevant for EMC Centera migrations.

Enterprise Vault requires a profile to have **Read**, **Write**, **Delete** and **Exist** capabilities in order to perform all its operations. Enterprise Vault version 5.0 or later checks that a connection has the Read, Write, Delete and Exist capabilities it needs.

Pool Entry Authorization (PEA) files

The use of Pool Entry Authorization (PEA) files – files containing authentication credentials for a given profile – is a best practice for applications when connecting to EMC Centera during the Centera Application Authentication process. The following are the Enterprise Vault requirements for this PEA File:

- A profile with **Read**, **Write**, **Delete** and **Exist** capabilities.
- The PEA file must be accessible to the Enterprise Vault Storage Service that manages the Vault Store Centera partition that will use the profile.

- The Storage Service runs under the Vault Service account and therefore the PEA file must allow read access by this account.
- For greater security, access to the PEA file by any other account could be disabled.

BEST PRACTICE: Usage of Pool Entry Authorization files is strongly recommended.

Note: If Application Authentication is not used – which is not recommended – an application is granted access to EMC Centera using the Anonymous profile; to prevent this, the Anonymous profile is disabled by default. Enabling this profile is only available to the EMC Centera system administrator.

Replication

For disaster recovery and high availability purposes, it is recommended to have at least EMC Centera unidirectional replication enabled; ideally, bidirectional replication should be set up.

Note: Failure to follow the standard procedures for pools and profiles creation and PEA files generation on replicated environments may lead to Enterprise Vault not being able to successfully establish an EMC Centera connection.

Details for these procedures are available at the EMC Centera online help.

Network segmentation

The concept of network segmentation – the use of multiple physical networks allowing each type of traffic to be segregated, monitored, and managed according to the appropriate per-site policies – for different types of traffic going to the EMC Centera system adds additional security for end users. Any management interaction with the system can be segregated from application input/output, ensuring the utmost security for data integrity

Note: For environments where such separation of traffic is not necessary, a single physical network may still be used for all network traffic.

The main characteristics of the different network traffic types going to and from EMC Centera are:

Application traffic

- Input/Output requests are defined as standard input/output operations to the EMC Centera (read, write, delete, exists, and query).
- Replication requests are defined as requests to a secondary or disaster recovery EMC Centera to store data.

Management traffic

- Centera Viewer/CLI connections are administrative tools for system administrators and end users.
- E-mail home mails are reports from the EMC Centera that show system health, failure alerts, and sensors.
- SNMP traps are additional monitoring of system health.

Node roles

The roles that can be assigned to each individual node are either external or internal:

- **External node roles**: Nodes with an external node role have an external IP address configured and use their Eth2 port for communication with the customer's network.

The external node roles are:

  - **Access**: For all application transactions. A node configured with the access role controls all input/output operations to and from the EMC Centera. The access node also serves as the staging area for all content that is archived and retrieved. The access nodes control all requests that come from application servers to store, retrieve, or dispose content.
  - **Replication**: For all replication and restore traffic between clusters.
  - **Management**: For all manageability connections to the cluster (CV, CLI, and Centera Console) and the E-mail Home, Syslog, and SNMP services.

- **Internal node roles**: Nodes with an internal node role do not allow for any communication with the customer network.

The internal node roles are:

  - **Storage role**: The storage role is an internal node role. Nodes with the storage role store data. It is possible to combine the storage role with one or more external node roles (not on Gen2 nodes).
  - **Spare nodes**: Nodes without any node role are spare nodes.

Note: Additional information for spare nodes is available in the Centera online help.

It is recommended that a thorough assessment is done both of the cluster when deciding which nodes should be assigned which roles and of the traffic expectations. The access and replication roles require the most system resources to complete their tasks. It is recommended that newer hardware (Gen4 and Gen4LP, if available) be used for these roles. Management traffic requires far less system resources and can be assigned to older generations of hardware (Gen2 and Gen3).

Application failover

Application failover is affected by network segmentation. The EMC Centera SDK has internal capabilities to retrieve data from multiple EMC Centera systems. This is achieved by leveraging the replicated configuration on the EMC Centera. When the application interacts with EMC Centera, the replication settings are stored by the SDK for purposes of failover. If a primary EMC Centera in a traditional failover environment becomes unavailable, the application still has the capability to retrieve data from the replicated environment.

Note: In a purely segregated environment of application and replication data traffic, built-in failover functionality will not function as expected. It is recommended that the access role is added to the nodes with replication role in the replica Centera.

Compliance Mode

EMC Centera compliance models

EMC Centera offers three compliance models or editions: Basic, Governance Edition (GE), and Compliance Edition Plus (CE+).

Note: EMC Centera relies solely on the application to perform any disposition actions, as data will never be deleted automatically/proactively; depending on the compliance model, EMC Centera will prevent the deletion of data still under retention.

Basic model

In its Basic edition, EMC Centera delivers the full power of content addressed storage (CAS). Self-configuring, self-managing, and self-healing, it captures and preserves original content, protecting the context and structure of electronic records. However, data retention is not enforced and advanced features such as shredding and advanced retention management are not available. Data can be deleted at any time, provided the application has the appropriate access rights, regardless of the retention period initially set.

Governance Edition (GE) model

Governance Edition provides the retention capabilities required by organizations to responsibly manage electronic records, on top of the features provided by EMC Centera Basic. Deploying Governance Edition enforces organizational and application policies for information retention and disposition. Original content can be captured and preserved —and ensure complete, reliable integrity for the life of the archived information.

Compliance Edition Plus (CE+) model

Compliance Edition Plus exploits the core strengths of the EMC Centera platform while adding extensive compliance capabilities to the Governance Edition model. CE+ is designed to meet the requirements of the most stringent of regulated business environments for electronic storage media as established by regulations from the U.S. Securities and Exchange Commission (SEC), the Australian AS 3806 Compliance Programs, or other national and international regulatory groups.

Note: Upgrading from Basic to either the GE or CE+ compliance model will have an immediate effect on retention reinforcement for all legacy data already archived.

The following hierarchical retention period setting applies to each data object stored on EMC Centera:

- **Application setting**: The application can assign a fixed retention period or a retention class to the CDF during its creation. The application setting of the retention period overrules the pool and cluster retention setting. Note that a CDF can be deleted only if the retention class to which it is assigned to is defined on the cluster.
- **Pool setting**: If a CDF does not have a retention period or class assigned by the application, the *default retention period of the pool* applies to the CDF. The pool setting of the retention period overrules the cluster retention setting.
- **Cluster setting**: If the CDF does not have a retention period or class assigned by the application and if the pool has no default retention period, the *default retention period of the cluster* applies to the CDF.

Retention periods

An EMC Centera retention period provides a simple mechanism for defining retention periods for items archived in Centera. EMC Centera retention periods are fixed numeric values part of the metadata (CDF) that cannot be changed; if a new retention period is required, a new metadata file (CDF) must be created for that piece of content. The data will be eligible for deletion only after all retention periods associated to it have expired.

Retention classes

Retention classes provide a way to manage and change retention periods for a set of data objects. Contrary to fixed retention periods given by the application or the pool, the retention periods assigned to a retention class can be changed by the system administrator. A retention class exists as a symbolic representation of a retention period.

A retention class name—not the period itself—is associated with a CDF. A retention class acts as a retention policy that governs all CDFs cluster-wide for those CDFs assigned to that retention class. If the time period of a retention class is changed, it likewise immediately affects the retention period of all CDFs referring to that class, without changing the individual CDFs. EMC Centera supports up to 1,000 retention classes, which can be defined only via the CLI. The SDK allows the setting up and removal of retention class assignments on CDFs. A retention class period can be increased and decreased at GE and only increased at CE+

Note: Retention periods in EMC Centera are calculated from the date that the CDF was created (archival or creation date) and not from the item's last modification date.

Enterprise Vault supports both EMC Centera retention periods and retention classes, referred as retention categories. It is possible to establish Enterprise Vault's retention behavior based on one of the following options:

- **Only for the EMC Centera Compliance Edition Plus model**: When selecting this option, Enterprise Vault sets a retention period on Centera Compliance Plus (CE+) models that corresponds to the Enterprise Vault retention period (defined in each item's retention category). On other EMC Centera models, Enterprise Vault sets a retention period of zero, which means that the items can be deleted immediately.

Note: If this option is set before the EMC Centera Compliance Edition Plus model is installed, Enterprise Vault begins setting the retention period only for all newly archived items that correspond to the Enterprise Vault retention period (all legacy data has a zero retention period).

- **For all Centera models**: When selecting this option, Enterprise Vault sets a retention period on all EMC Centera models that corresponds to the Enterprise Vault retention period (defined in each item's retention category).

- **Never**: Select this to make Enterprise Vault set a retention period of zero on all EMC Centera models (items can be deleted immediately).

When configuring retention policies in Enterprise Vault, it is recommended to follow the guidelines below, depending on the nature of the data to be archived:

- **Retention classes**: for any new data eligible for archival

- **Retention periods**: for any legacy or backlog data to be archived

Note: Enterprise Vault retention categories should be mapped to EMC Centera retention classes during the setting of the Vault Store retention policies.

It is the role of the application to set the retention periods that EMC Centera enforces, and to store, retrieve, and dispose of (delete) content as required. Once the retention period of a content object has expired, the application must dispose of the content by way of the Centera API; normal delete operations will fail on GE and CE+ models when attempting to dispose objects still under retention. A cluster will never proactively dispose of content managed by an application.

Although Enterprise Vault does not explicitly recommend data deletion for content still under retention (commonly known as privileged deletes), users with enough access rights are allowed to perform such activities. For implementations where EMC Centera has been deployed using GE or CE+ compliance modes, these actions might result in warning or error messages reported in the Administration Console, as EMC Centera will deny such requests until the retention period is expired.

As Enterprise Vault does not leverage EMC Centera Privileged Deletes, when the delete request is submitted, Enterprise Vault will attempt and fail the delete operation, and the end user will be notified. However, if the Expiry Service attempted to delete the item, Enterprise Vault will try again the next time the service runs; this operation will be retried each time the Expiry Service runs until the retention period expires and Centera grants the delete request.

It is recommended to carefully review all business cases before defining Enterprise Vault retention policies and the roles/access rights that users will have, in order to prevent such temporary processing overheads and data inconsistencies. (For additional information, see the white paper *EMC Centera Compliance Models – A Detailed Review)*.

Installation process

Licensing

Enterprise Vault uses the Enterprise Licensing System (ELS). To run the associated Enterprise Vault Services, a license key file that covers the Enterprise Vault features that the user wants to implement must be installed.

The following types of Enterprise Vault license are available:

- **Production license.** This license comprises a product base license and any additional feature licenses. When the license file is installed, the functionality of Enterprise Vault depends on the feature licenses that were purchased. Production licenses generally do not have an expiry date.

- **Trialware license.** With this 30-day license, the full functionality of Enterprise Vault is available, but the functionality is time-limited, as defined by the key. When the license expires, the software continues to run in restricted, read-only mode, which allows archived items to be viewed and retrieved, but no items can be archived. Enterprise Vault tasks will not start, and the contents of personal folder (PST) files cannot be migrated to Enterprise Vault.

- **Temporary licenses.** Temporary licenses are available for 10-day to 90-day duration.

When the license expires, the software continues to run in restricted, read-only mode, which allows archived items to be viewed and retrieved, but no items can be archived. Enterprise Vault tasks will not start, and the contents of PST files cannot be migrated to Enterprise Vault. For additional information, see the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

Installing Enterprise Vault

The following are the steps needed to install the required Enterprise Vault components:

1. Log in to the Vault Service account to install Enterprise Vault.
2. Load the Enterprise Vault media.
3. Open the Symantec Enterprise Vault 8.0 folder.
4. Double-click the ReadMeFirst file to display the ReadMe text and read it before continuing with the installation.
5. Open the Server folder.
6. Double-click SETUP.EXE to start the installation.
7. Install the required Enterprise Vault components for this computer.

The core components for an Enterprise Vault server are as follows:

- Enterprise Vault Services - Installs the entire core Enterprise Vault Services. After the installation, the services must be configured before using them. This is done when the Enterprise Vault configuration wizard is run. (Additional information is available in the "About configuring Enterprise Vault" section of the *Symantec Enterprise Vault 8.0 – Installing and Configuring Guide*).

- Administration Console - Installs the Administration Console. This is a snap-in to the Microsoft Management Console (MMC) that enables the user to manage Enterprise Vault.

  This component also installs the Enterprise Vault configuration wizard, PST Migrator, and NSF Migrator wizards.

  To install a standalone Administration Console on a remote system, select this component only.

  A number of other components can be installed as required, if their prerequisites are met. Some of these components are listed only if certain software is present:

  – SMTP Archiving Components, Exchange Server Extensions, and Microsoft SharePoint components are usually installed on computers other than the Enterprise Vault server. For details, see the appropriate section of the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

  – Enterprise Vault Operations Manager must be installed on at least one Enterprise Vault server if the desire is to use it to monitor the Enterprise Vault servers at that site.

  – Enterprise Vault Reporting is listed for selection only if Microsoft SQL Server Reporting Services (SSRS) are installed on the computer. Enterprise Vault Reporting can be installed on an Enterprise Vault server, but is more typically installed on a separate server running SSRS.

8. At the end of installation, a computer restart might be required.

Post-installation tasks

Default security for the Web Access application

The default security settings for the Web Access application configure automatically, during the Enterprise Vault installation, and set access to the Web Access application using HTTP over TCP port 80; both Basic authentication and Integrated Windows authentication are configured automatically.

These settings affect the way users log in to the Web Access application, as follows:

- For Internet browsers supporting Integrated Windows Authentication (for example, Internet Explorer)

  The user must supply domain name and username separately:

  - Username: username
  - Password: password
  - Domain: domain (This domain can never be defaulted)

Note: An Internet Explorer user with suitably customized browser settings does not need to supply login details manually because the login is automatic; Internet Explorer automatically uses the details of the account to which the user is currently logged in.

See "Customizing security on the client computers" in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

- For Internet browsers that do not support Integrated Windows Authentication the user must supply both domain name and username in response to a single username prompt:

  - Username: domain\username
  - Password: password

It is possible to set up a default domain. See the section "Customizing authentication and Customizing security for Web Access," in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

Note: If a message was received during the installation saying that setup could not set alias security, please refer to the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* "Setting up the default authentication" section, for detailed instructions on how to perform the default authentication.

Customizing security for the Web Access application

In addition to customizing the amount of information that users need to provide when logging in to the Web Access application, it is also possible to change the port or protocol that is used to access the Web Access application, if for instance it is required that connections to the application be made using HTTPS.

The *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* has additional information on how to perform these activities.

Customizing security on the client computers

It is possible to configure Internet Explorer on user computers so that users are automatically logged in to the Web Access application, without receiving a login prompt. Essentially, Internet Explorer must be configured so that it trusts the Web Access application computer.

For this to work, the Integrated Windows Authentication is a requirement, as described in "Setting up the default authentication" in the *Symantec Enterprise Vault 8.0 – Installing and Configuring Guide.*

To make Internet Explorer log in automatically, it may be required to modify the Internet Explorer Internet Options on each client computer. As the settings are saved in the Windows registry, it is possible to save them for en-mass rollout purposes.

Out of the many ways available to configure Internet Explorer security, the methods listed are detailed in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide*

- Using the proxy bypass list
- Explicitly naming the Web Access application computer

Configuring Enterprise Vault

The Enterprise Vault configuration wizard can be executed either immediately after installation (after restarting the computer if prompted), or after performing the post-installation tasks for the Web Access application, as described in the previous section.

Note the following:

- If the configuration wizard is run immediately after the installation, remember that there are some additional tasks that need to be done before users can use Enterprise Vault.
- If the user exits from the configuration wizard before configuration is complete, the configuration wizard can be run again and have the option to delete the Directory database. Once the configuration wizard has been successfully completed, it cannot be run again on the same computer.

Using the Enterprise Vault configuration wizard it is possible to:

- Select which SQL server to use for the Enterprise Vault Directory database
- Create the Enterprise Vault Directory database (*)
- Create the Enterprise Vault Monitoring database
- Create an Enterprise Vault Site (*)
- Add the computer to the site (*)
- Select the Enterprise Vault Services that are desired to run on the computer
- Choose the storage areas to use for Enterprise Vault data

Note: Tasks identified with (*) can only be performed using the configuration wizard; however, other tasks such as adding a service or assigning storage areas for the data, can also be done using the Enterprise Vault Administration Console.

The main activities during the configuration process are:

- Registering the Enterprise Vault Service account on the local computer
- Creating the Directory Database (for the first Enterprise Vault server only)
- Creating the Enterprise Vault Site and registering the local computer
- Selecting the Enterprise Vault Services to be added to the computer
- Defining the storage locations for the Indexing and Shopping Services
- Configuring the Service mailbox
- Starting the selected services on the local computer
- (Optional) Launching the Getting Started wizard or the Vault Administration Console (VAC) to set up archiving.

Note: These instructions apply to a non-clustered environment. If Enterprise Vault is being configured in a Veritas Cluster Server or Microsoft Server Cluster environment, see the appropriate clustering section in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* .

Complete details on the Configuration wizard are available in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* and a summary of these tasks is presented elsewhere in this document.

The configuration wizard may be launched after completing the Installation Program or after restarting the computer, if it was initially required.

The first task to be executed is to register the Enterprise Vault Service account to the Local Administrators Group with the following advance User Rights:

- Log On As a Service and Act as Part of the Operating System
- Debug programs
- Replace a process-level token

When running the configuration wizard on the site's first Enterprise Vault server, a Directory database will be created on the specified SQL server (or instance); the location of the SQL server, as well as the path for the transaction log and database files, can be specified/modified at this time.

With the Directory database created, the configuration wizard will then create an Enterprise Vault Site and will add the local computer to the site using an unqualified DNS alias, as fully-qualified DNS aliases are not recommended.

Once the Enterprise Vault Services to be executed in the computer are selected and added to the computer, the default storage locations for the Indexing and Storage Services are defined; it is recommended to ensure that these locations are on an accessible device and the Enterprise Vault Service account has Write access to them. For Exchange archiving, special attention must be placed to these locations as they cannot be easily modified after they have been enabled. Before

mailboxes are enabled, the Administration Console can be used to change the index storage location, or add further locations. Then, the Service mailbox is configured, the selected services are started and their status are presented; selecting **Next** on the configuration wizard will refresh the status, until all of them are successfully started. It is recommended to check the licenses keys to ensure the selected services are properly licensed.

The final screen of the wizard presents the following options:

- Run the Enterprise Vault Getting Started wizard - It is recommended to select this option, as it allows setting up archiving as quickly as possible. The wizard provides both express and custom options for maximum flexibility (for additional information, see *Setting up Archiving*).
- Run the Enterprise Vault Administration Console - This option is only recommended for users already familiar with the Administration Console and familiar with setting up archiving.
- Just close the wizard - This option will close the Configuration wizard. Access to the Enterprise Vault Getting Started wizard or the Administration Console is available via the Enterprise Vault Start menu options.

It is also recommended to refer to the Enterprise Vault online help when using the configuration wizard to configure Enterprise Vault on subsequent computers.

Setting up archiving

Up to this point, all information provided in this document has been generic to the installation and configuration of Enterprise Vault. It is here when the best practices for archiving to EMC Centera come into play, as the storage device (EMC Centera) is selected while setting up archiving.

EMC Centera configuration

Prior to performing the Enterprise Vault archiving setup, EMC Centera is expected to be fully configured; in order to ensure Enterprise Vault access to Centera, at the end of this process, the following items must be available to the Enterprise Vault system administrator:

- Generated PEA file
- Source (primary) Centera access node(s) IP address(es)
- If replication is enabled on a network segmented environment, target (replica) Centera access node(s) IP address(es)
- Name of default Centera retention class (optional)

Note: If replication is enabled on an environment being replicated without leveraging Centera network segmentation capabilities, the IP addresses of the replica cluster are not required.

The main activities during the Centera configuration process are:

- Setting the initial configuration according to the EMC Change Control Request Form (CCRF) (factory defaults CPM and SSP recommended)
- Creating the Enterprise Vault application pool
- Creating the Enterprise Vault application profile

- (Optional) Setting IP restrictions on the application profile
- Generating the associated PEA file on the primary Centera cluster
- Creating the default retention class
- Setting up replication
- Generating the associated PEA file on the replica Centera cluster
- Merging primary and replica PEA files

Complete details on these procedures are available in the EMC Centera online help, and all of them are performed by certified EMC Implementation Specialists using the Centera Viewer / Command Line Interface (CV/CLI).

The first task to be executed is the configuration of the cluster according to the guidelines laid out on the CCRF; the particular importance are the settings for Content Protection, which should be set to Mirroring (CPM) and the Storage Strategy, to be set to Performance Full (SSP).

When creating the application pool, the **Write**, **Read**, **Delete**, and **Exist** capabilities must be defined as the Pool Mask; any rights omitted will deny that particular capability. Any requirements for pool quota management are specified at this time.

Next, an application profile with Data Access Capabilities must be created, having the application pool created as the home pool. Answering "Yes" when prompted to establish a Pool Entry Authorization will allow the definition of the location of the PEA file in the local machine where the CV/CLI utility is installed.

Note: This location is temporary and only used to generate the file; once created, the file must be located on a shared path accessible to the Enterprise Vault Storage Service (see Access Nodes Addresses for more details). Generation of the PEA file is possible at any other time by the use of the "update profile" CLI command.

Optionally, IP restrictions on the profile to only allow the application and the management servers access to the pool can be specified.

To finish the setup on the primary cluster, the creation of the necessary retention classes is performed. An EMC Centera retention class must be created to be used as default, and should mirror the retention policy of the Enterprise Vault default retention category.

Note: For security reasons, retention classes in EMC Centera can only be created by an EMC Centera system administrator; it is recommended that mapping and consistency between Centera retention classes and Enterprise Vault retention categories is ensured by the appropriate system administrators.

Note: New retention classes can be defined at any time, based on business requirements. Updates to existing retention classes are possible, depending on the Centera Compliance mode in use.

If replication is being leveraged for disaster recovery and high availability purposes, the pool and profile configurations must be exported from the primary EMC Centera cluster and imported into the replica cluster; similarly, all retention categories must be created. If network segmentation has been implemented, the Access Role must be added to the nodes with the Replication Role on the target cluster.

Once the replica cluster is set up, its PEA file must be generated, and both PEA files are merged into a unique file to be used for any and all connection purposes. Detailed information on merging PEA files is available in the EMC Centera online help.

Enterprise Vault configuration

Note: Prior to setting archiving, access to the Centera PEA file, and the IP addresses of the access nodes on the primary and replica addresses (if replication is enabled) is required. See Centera Configuration for additional information.

Although archiving can be manually set up using the Vault Administration Console (VAC), it is recommended to use the Getting Started wizard in order to set up archiving as quickly as possible.

The wizard can be executed immediately after the Configuration wizard as part of a new Enterprise Vault deployment, or from the Enterprise Vault Start menu options. The Getting Started wizard cannot be executed in the same computer once it has successfully finished; however, it can be used on other computers in the site, or restarted in the same computer if it was interrupted before the end.

The Enterprise Vault Getting Started wizard assists in the following tasks, as appropriate:

- Set up storage locations
- Create retention categories
- Create archiving policies for Exchange Server, Domino, and File System Archiving

It is possible to run sections of the wizard in express mode or in custom mode, as follows:

- In express mode, the wizard does not ask many questions. Instead, it applies as many default settings as possible. Later, the Administration Console can be used to make changes to the settings, if required.
- In custom mode, the wizard provides the flexibility to change the default settings.

Note: It is required to run the wizard in custom mode for the storage configuration section in order to configure EMC Centera as the remote storage location – which by default is set to a local NTFS volume that cannot be modified after the fact – for the Vault Store Partition's Device type (EMC Centera).

It is required that the Enterprise Vault license keys be installed as the Getting Started wizard checks them to determine which options to present. It is recommended to run the Enterprise Vault Deployment Scanner to determine whether the Enterprise Vault prerequisite configuration is correct.

Additional information is available in the Deployment Scanner manual in the Documentation folder of the EnterpriseVault media.

Storage configuration

Storage in Enterprise Vault is logically allocated to each Vault Store Partition; a Vault Store Partition is part of the Site hierarchy, set as the lower level within the Vault Store Group/Vault Store branch. For EMC Centera, the Getting Started wizard will set up the Enterprise Vault components listed in Table 9.

Note: Refer to the "Planning for the Getting Started Wizard" chapter in the *Symantec Enterprise Vault 8.0 - Installing and Configuring Guide* for a complete planning sheet listing the Getting Started wizard's Express-mode choices:

Table 9 Enterprise Vault components – Summary of recommended values for EMC Centera devices

Component	Item	Recommended EMC Centera value
Vault Store Group	Not Applicable	Although the creation of a Vault Store Group and a Fingerprint database are required for Enterprise Vault 8.0, if only EMC Centera partitions are used, none of the settings on the Vault Store Group are relevant; therefore, all values can be left at their default, for example, Sharing. The Fingerprint database is created but not used; no special care needs to be taken when deciding where to place the database and log files on disk.
Vault Store	Sharing	Not applicable. For EMC Centera the setting is defined at the Vault Store Partition level.
Vault Store	Remove Safety Copy	After backup
Vault Store Partition	Storage Type	EMC Centera
Vault Store Partition	Access Nodes Addresses	The network fixed IP addresses assigned to the EMC Centera access nodes
Vault Store Partition	Sharing	"Enable device-level sharing" ON
Vault Store Partition	Retention Period	For all EMC Centera models
Vault Store Partition	Check Centera Replication (interval)	60 minutes
Vault Store Partition	Enable collections	"Enable collections" ON
Vault Store Partition	Location for Temporary Files	For collections, the temporary location where the collections will be created (at least 1% of estimated Centera capacity; 50 GB recommended)
Vault Store Partition	Security ACLs	Checked

Note: The wizard will first go through the process of creating a Vault Store Group, as a requirement for Enterprise Vault 8.0, although it is not relevant when EMC Centera is the only type of partition used. Similarly, when creating a new Vault Store, the Sharing tab is not relevant and is ignored for Centera partitions, as this will be defined at the Vault Store Partition level instead.

Remove safety copy

Enterprise Vault can be configured to retain archived items until the Vault Store partition in which they are archived has been backed up. During the time between archiving and removal, the original items are treated as safety copies by Enterprise Vault.

This feature favors compliance requirements around the need for keeping two copies of the same content available at all times; similarly, this could be leveraged during disaster recovery procedures.

When the Vault Store partition has been backed up, Enterprise Vault can remove the safety copies. It also creates shortcuts and placeholders at this time if it is configured to do so. See the

"Managing Safety Copies" section in the "Day-to-day Administration" chapter of the *Symantec Enterprise Vault 8.0 - Administrators Guide* for additional information.

When deployed on an EMC Centera replicated environment, Enterprise Vault's archived data is replicated instead of backed up, significantly reducing the backup window requirements, as well as providing for a high availability environment.

It is recommended to set the removal of the Safety Copies After backup, so that Enterprise Vault ensures the existence of the replicated data before removing the safety copies from primary storage (for example, MS Exchange) (see Figure 4).

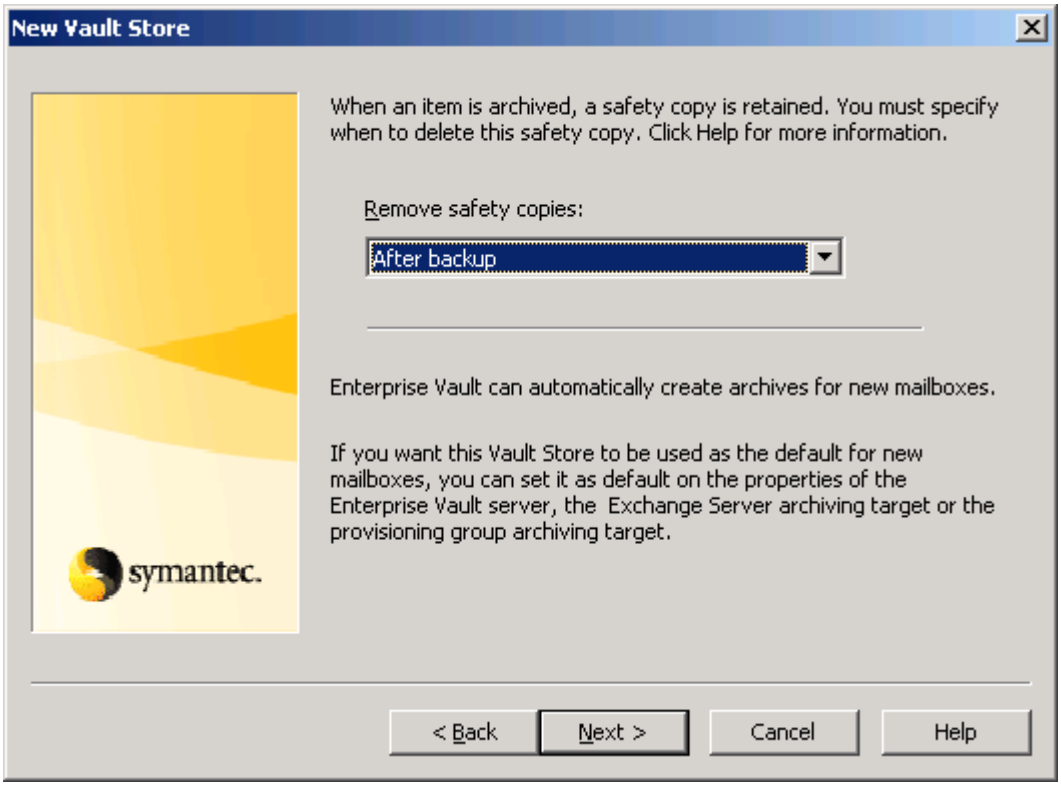

Figure 4 Setting up removal of safety copies during vault store creation

The removal of safety copies can also be configured, using the Vault Store Properties page.

To configure the removal of safety copies for an existing Vault Store, go to the General tab of the Vault Store Properties page, and choose the **After backup** option.

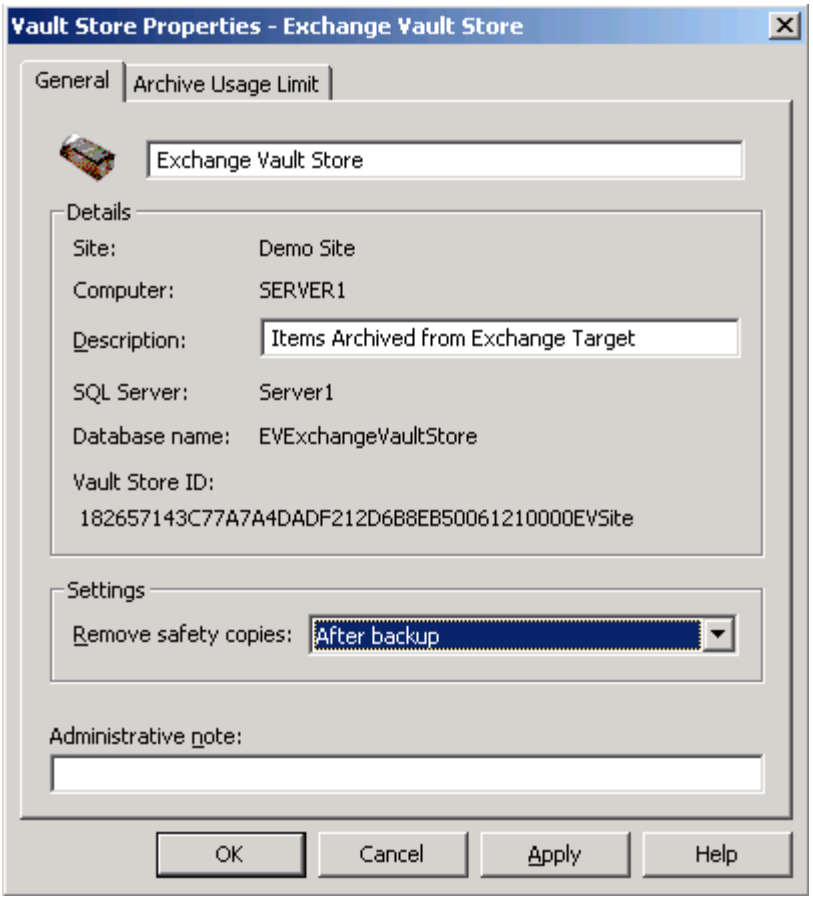

Figure 5 **Setting up removal of safety copies for an existing Vault Store**

Note: For fine-tuning purposes, the Remove Safety Copies option could be set to After Backup (Immediate for Journaling); for journaling, this means that the existence of the items in the replica Centera is not verified, to save on the performance hit in EV of post-processing the items.

Storage type

The wizard will then go through the process of creating a partition. Selecting EMC Centera as the storage type will start a configuration path different from other partition types (Figure 6).

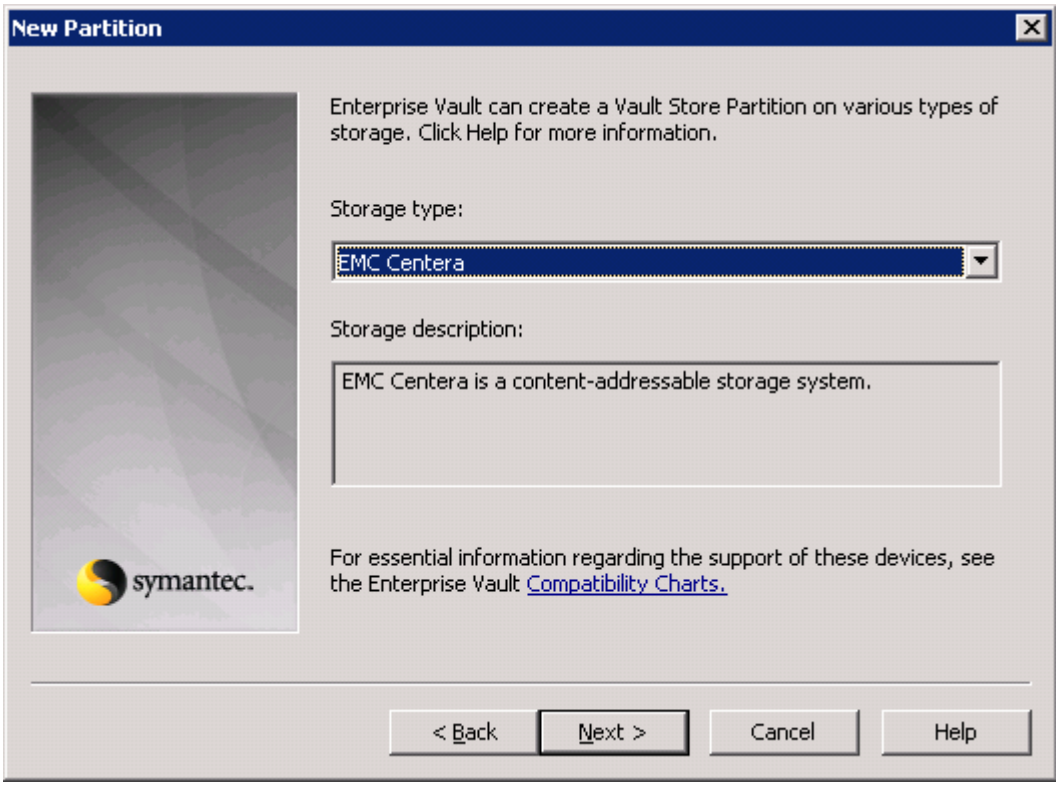

Figure 6 **Selecting Centera as the storage type**

Note: Additional information regarding Symantec's compatibility and support for EMC Centera and other devices can be found on the Enterprise Vault Compatibility Chart.

Access node addresses

Next, access node IP addresses are entered (Figure 7) by clicking **Add** to add a new IP address, until all access node (AN) IP addresses are registered.

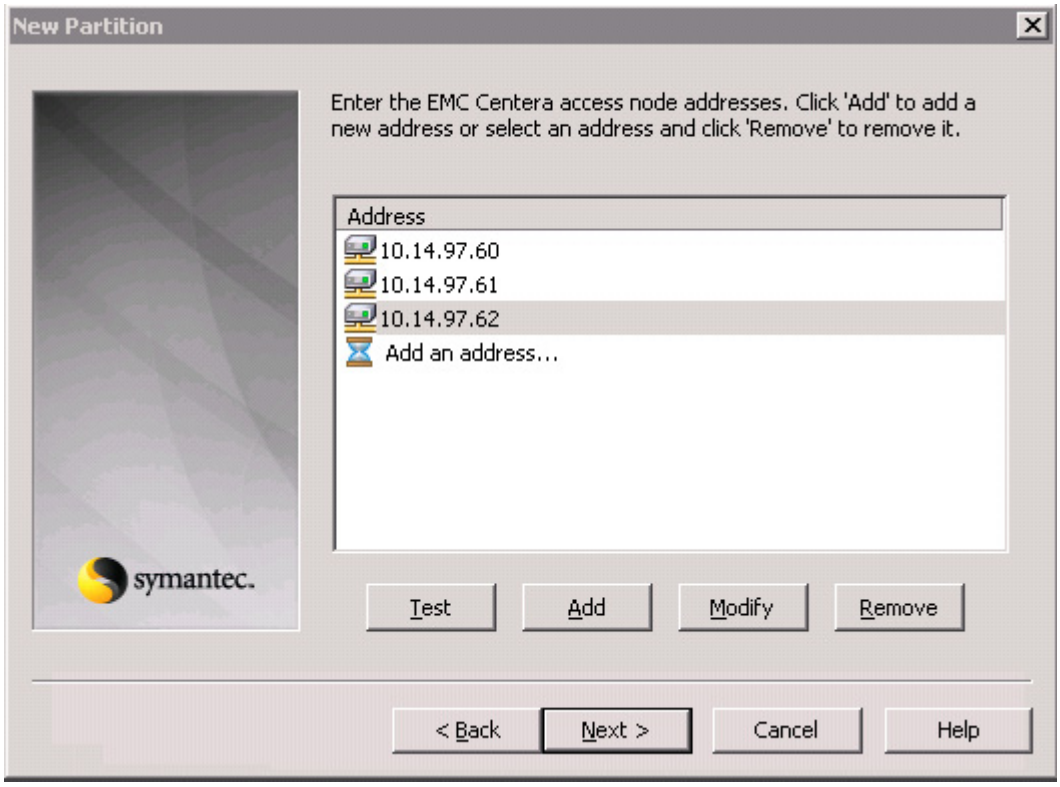

Figure 7 **Entering the EMC Centera access node (AN) IP addresses**

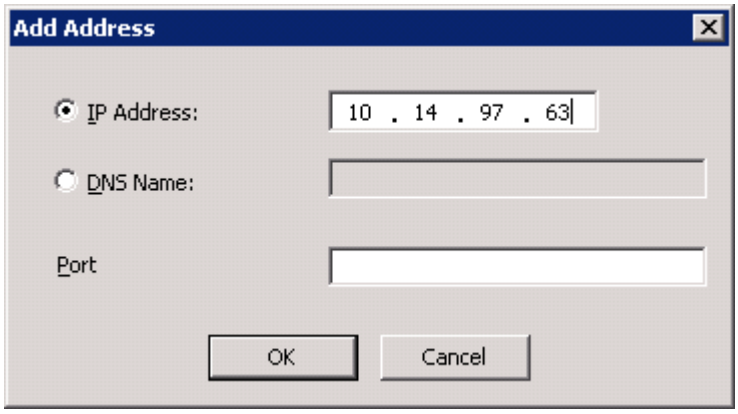

Figure 8 **Add Address (DNS Name optional)**

Note: Optionally, defining a DNS Name for each access node will help minimize any impact related to future infrastructure network changes on Enterprise Vault's access to EMC Centera.

Sharing

EMC Centera single instancing (SiS) is enabled when setting the Enterprise Vault sharing feature by selecting the **Enable device-level sharing** checkbox (Figure 9).

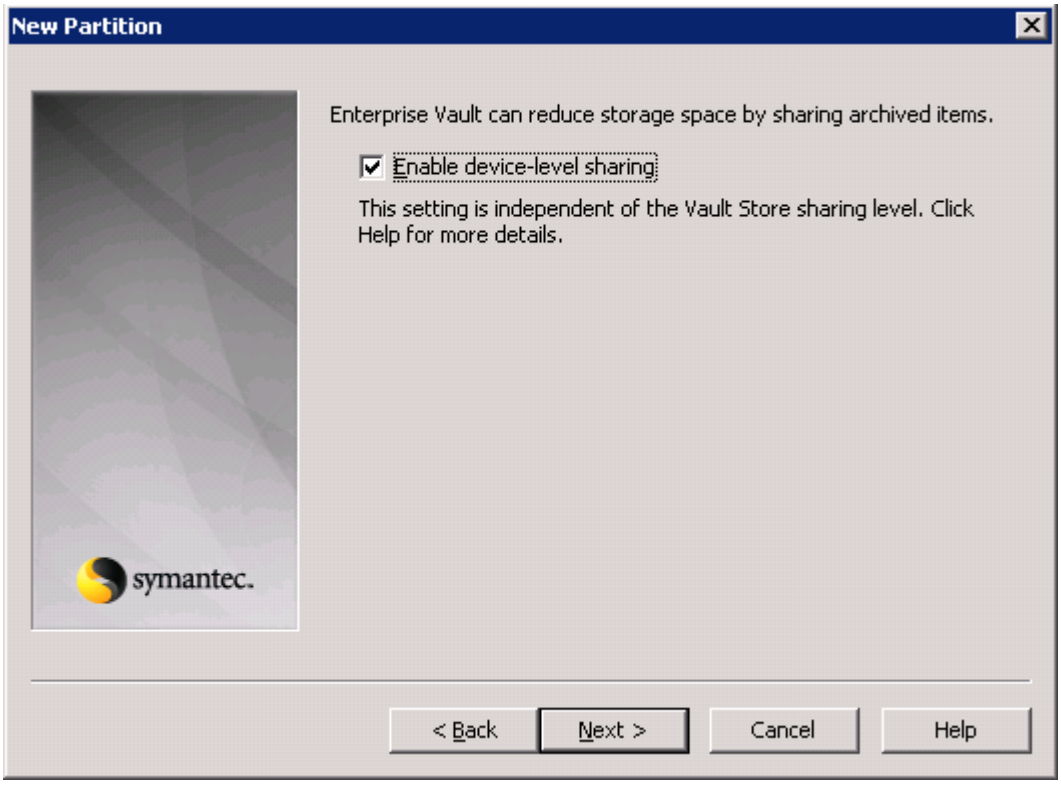

Figure 9 Enabling Centera SiS on the Enterprise Vault partition

The way that items are shared or single-instanced with EMC Centera differs from other devices. On EMC Centera, attachments are detached from the message and stored in Centera, where Centera identifies them as candidates for sharing. The rules are as follows:

- A saveset with an uncompressed size of 100 KB is stored unshared.
- A saveset with a compressed size of over 100 KB is examined for "streams"—indexable items or XML streams such as recipient lists—and attachments.

Note: In EV 8.0 and later the single "Indexable item" stream has been replaced with one or more "Convertible Content" streams.

- If there are no streams or attachments, the saveset is stored unshared.
- If there are no streams or attachments with an uncompressed size of over 50 KB, the saveset is stored unshared.
- Any stream or attachment with an uncompressed size of over 50 KB is stored separately and is eligible for sharing.

This model has the advantage that attachments are shared even if they are attached to different messages or archived separately by File System Archiving. It also means that there is sharing across Vault Stores. Small messages are not shared. However, even though small messages make up the bulk of messages, messages with large shareable attachments usually make up the bulk of the size. For example, a large report might be sent or forwarded to all members of a company. Just one copy of this report is held on EMC Centera, although there will be many copies held on the Exchange Stores or Lotus mail files in the company.

Enterprise Vault can optimize the use of storage space by storing a single instance of items that have copies in multiple places; for example, a large PowerPoint presentation sent to multiple recipients. To enable single instance storage, **Enable device-level sharing** must be enabled on the partition properties; this will ensure that EMC Centera single instancing is possible for multiple Vault Stores, Vault Store Groups, and even Vault Sites.

It is also possible to enable device-level sharing from the General tab of the partition properties (Figure 10).

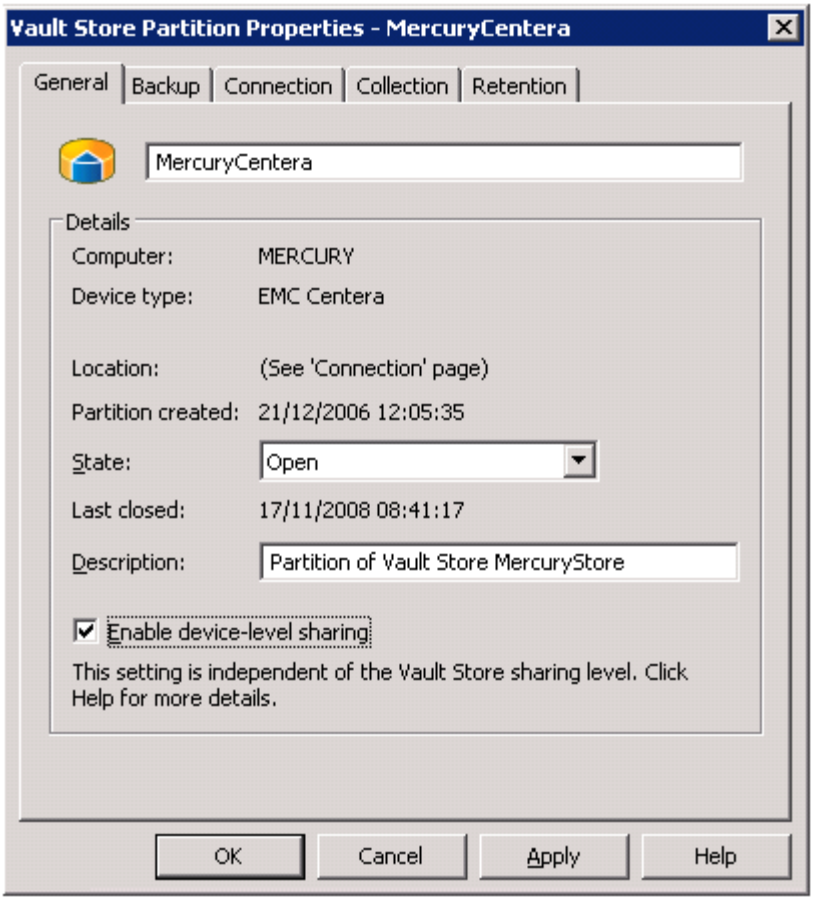

Figure 10 Setting up single instancing (sharing)

The Vault Store Partition Properties dialog box is displayed for the partition clicked. The General tab allows you to turn on sharing and set the state of the partition.

Note: Enabling Enterprise Vault device-level sharing is how attachments are written as single instanced BLOBs on EMC Centera. Failure to select this option will result in no SiS savings from EMC Centera.

Note: It is recommended to always have this checked in order to achieve EMC Centera SiS.

Note: For storage other than EMC Centera, Enterprise Vault SiSO capabilities must be carefully assessed as there are additional considerations on network traffic and extra complexity to the solution, such as fingerprint databases, policies, and limitations on data moves/migrations. See "Developing a suitable sharing regime" in the Enterprise Vault 8.0 Install Guide.

Note: Partitions for EMC Centera do not take part in Enterprise Vault single instance storage sharing. If a partition is created for EMC Centera in a Vault Store that is configured for sharing, the partition is ignored for the purposes of Enterprise Vault single instance storage sharing.

Retention period

To determine how the Enterprise Vault retention policy settings should be reflected in EMC Centera, the wizard presents three alternatives (see the "Compliance Mode" section). To have the system apply, at minimum, the default Enterprise Vault retention period for all data archived to EMC Centera, select the **For all Centera models** option.

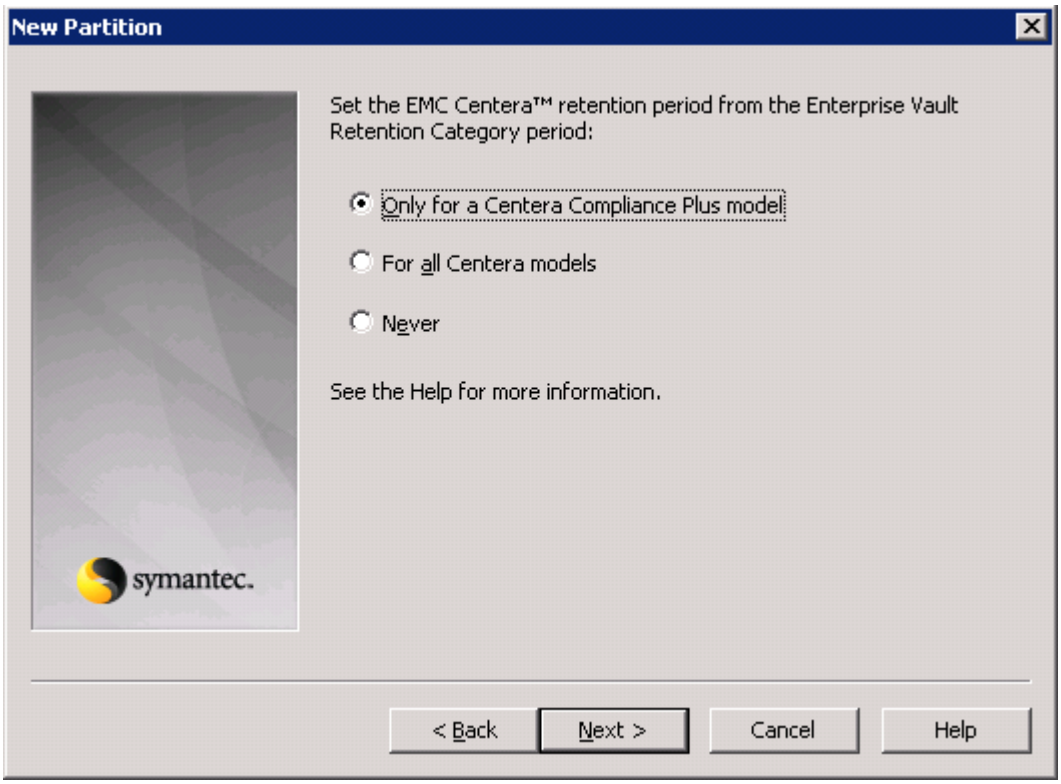

Figure 11 Setting retention policies from Enterprise Vault to Centera

Check Centera Replication

Setting up how often replication is checked (see Figure 12) supplements the Safety Copy configuration defined at the Vault Store level. This step is performed to determine the maximum time elapsed before verifying they have been replicated to the secondary Centera and thus becoming eligible for deletion from the Exchange server. It is recommended to take the default

60 minutes value as a fair compromise/balance between application performance and disaster recovery capabilities.

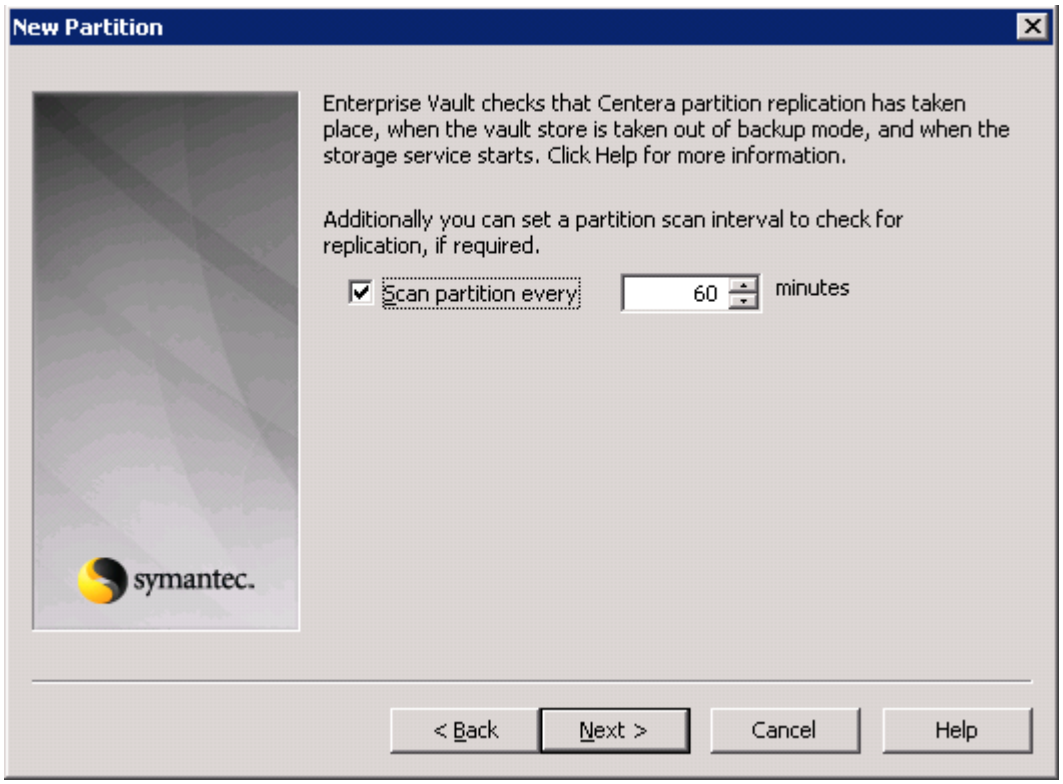

Figure 12 Defining Centera Replication's partition scan intervals

Enable collections

Enabling collections is the recommended value; the wizard will also prompt for the location of the staging area used to create/host the containers prior to sending to EMC Centera for archival (see Figure 13).

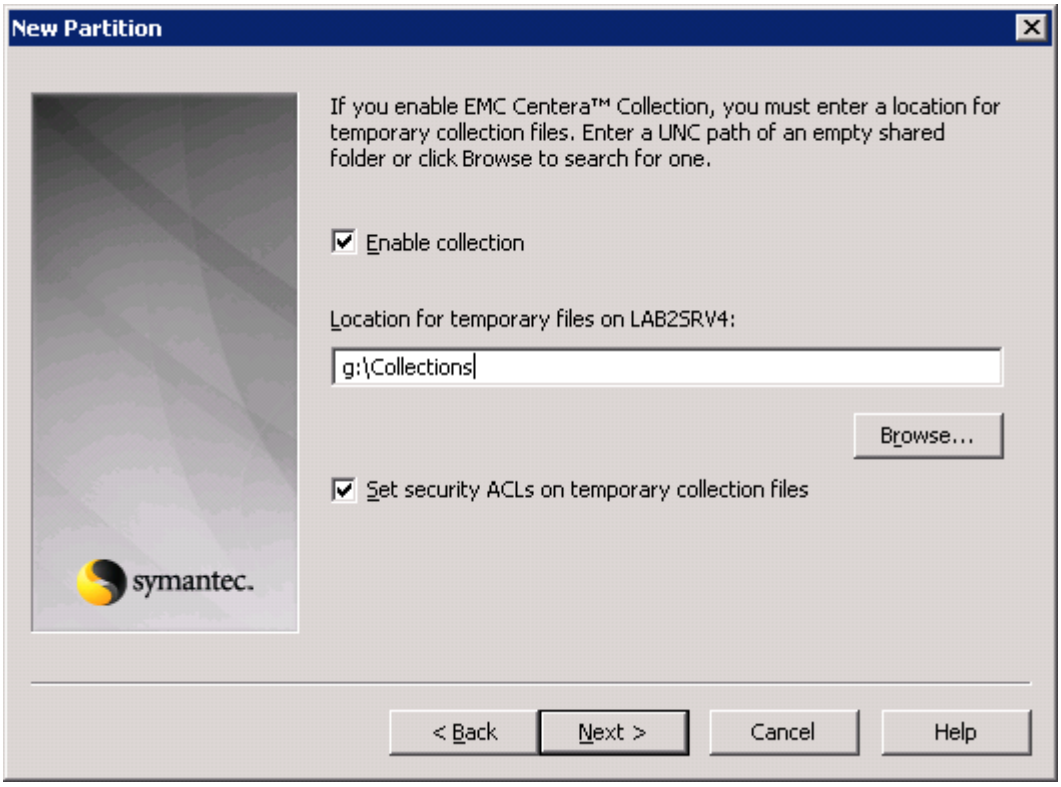

Figure 13 Setting up Centera collections

Enterprise Vault offers two methods of storing items in EMC Centera; with collections and without collections. EMC Centera collections are completely different from NT File System (NTFS) collections that can be used when storing to NTFS storage.

With collections, items are first archived to a local staging area. Another Enterprise Vault process collects files in this area and stores them on EMC Centera. There is some extra processing involved and CPU usage increases on the Enterprise Vault server. Despite this, similar archiving rates are usually achieved as when collections are turned off and there is no difference in retrieval rates.

A collection is up to 100 items or 10 MB of data. Collections are recommended because they result in fewer objects on EMC Centera. This has several advantages:

- No fall-off in performance as the EMC Centera gets fuller
- Fewer resources are used on the EMC Centera, allowing an overall greater throughput
- Faster replication
- Faster deletion of expired items

- Faster self-healing in the event of a failed disk
- Very fast retrieval of items because only the item is retrieved from EMC Centera and not the whole collection (Centera partial read)
- No impact on EMC Centera single instancing abilities

Additional information is available in the *Symantec Enterprise Vault 8.0 Performance Guide*.

Note: It is possible to see the collection configuration for existing Centera Vault Store Partitions, by selecting the Collection tab on the Properties dialog box (see Figure 14).

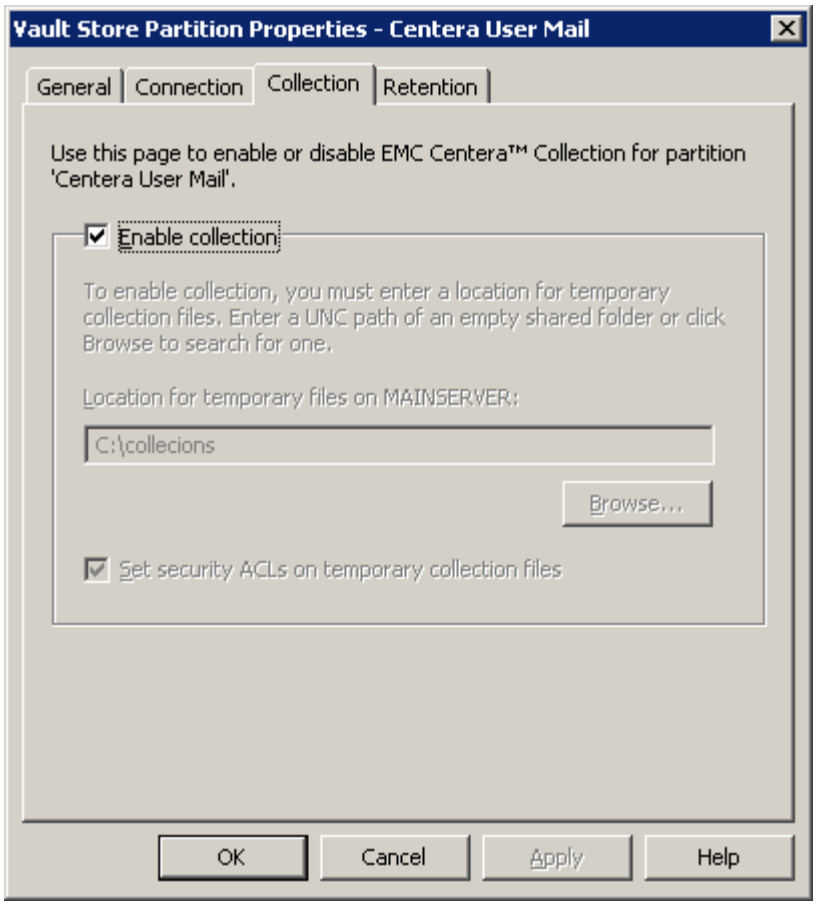

Figure 14 Enabling Centera collections on an Existing Vault Store partition

Collections and single instancing

Collections and single instance storage are handled differently on EMC Centera devices:

- Temporary saveset files are used instead of CAB files (used for other storage platforms).
- Files are collected as soon as they are archived (not according to a schedule).

Note: If Enterprise Vault Sharing (collections) is enabled, EMC Centera single instancing capabilities are not impacted.

When EV creates a collection, one data BLOB contains multiple objects. For each object an Attributes Tag and one or more Tags with data BLOBs for the separately stored Indexable Item Stream, XML Stream, and Attachments are created.

Note: In EV 8.0 and later the single "Indexable item" stream has been replaced with one or more "Convertible Content" streams.

These data BLOBs are capable of being shared and all other data BLOBs contain single attachments (or other streams) so that the attachments or streams can be shared (single instanced).

Note: Although archiving without collections is not recommended, if it has been decided to archive without using collections, then the number of processes writing to EMC Centera must be increased. Writes to ECM Centera take longer than to other devices, but many writes can take place in parallel.

Examples of increasing the number of processes are:

- To increase the number of Storage Archive processes, for Exchange Server mailbox and Exchange Server journal mailbox archiving, change the number of Archive Processes. Do this in the Administration Console by editing the properties of the Enterprise Vault Storage Service. Increase the number of Archive Processes to 10.
- To increase the number of Domino threads, from the Administration console, edit the properties of the Lotus Domino Task. Increase Number of concurrent connections to the Domino Server to 15.
- For File System Archiving and SharePoint the number of threads writing to EMC Centera should be 10. Change the value, if necessary, by editing the configuration files. There are examples of configuration files in the Enterprise Vault program folder, usually C:\Program Files\Enterprise Vault. These files are:
  - Example EvFileArcSvr.exe.config
  - Example EvSharePointArchiveTask.exe.config
- To change the number of threads, rename the file so that the name does not begin with Example, then edit the file, changing the value of **NoItemProcessorThreads** to 10.
  - 1. Find the following line: <add key="NoItemProcessorThreads" value = "10"/>
  - 2. Change the number to the new value as required.

Note: For additional information, see the "Threading" section.

Location of temporary files

Items for collection are stored on a local disk before they are archived to EMC Centera. This needs to be a fast disk but not large (See Local Disks and Requirements).

The temporary folder is where the collections are to be temporarily staged.

Security ACLs

It is recommended to use access control lists (ACLs) to further secure access to the collections.

Exchange Server archiving policies

When configuring Exchange Server targets, the Getting Started wizard searches the network for instances of Exchange Server. It is possible to select the Exchange Server computers for which archiving are to be configured.

For the selected Exchange Server the following must be determined:

- Specify whether to configure mailbox archiving or journal archiving, or both.
- For mailbox archiving, a system mailbox on that server that Enterprise Vault can use to log in must be specified.
- For journal archiving, the journal mailboxes to archive and the journal archive to use for each mailbox must be specified.

  The wizard enables the creation of new archives, if required.

The wizard will set up the Enterprise Vault components. Table 10 lists the items related to EMC Centera.

Table 10 Recommended default retention category for EMC Centera

Component	Item	Recommended EMC Centera value
Exchange Provisioning Group	Default retention category	Mapped to default retention class in EMC Centera, according to policies

BEST PRACTICE: It is recommended to create an EMC Centera retention class to map the Enterprise Vault default retention category.

If Enterprise Vault is configured to archive from the Exchange managed folders, it can automatically synchronize managed content settings to managed folder retention categories. Enterprise Vault creates managed folder retention categories automatically. For more information, see the *Symantec Enterprise Vault 8.0 - Administrators Guide*

When setting the retention policies, the user is required to associate the Enterprise Vault retention categories to the EMC Centera retention classes (Figure 15).

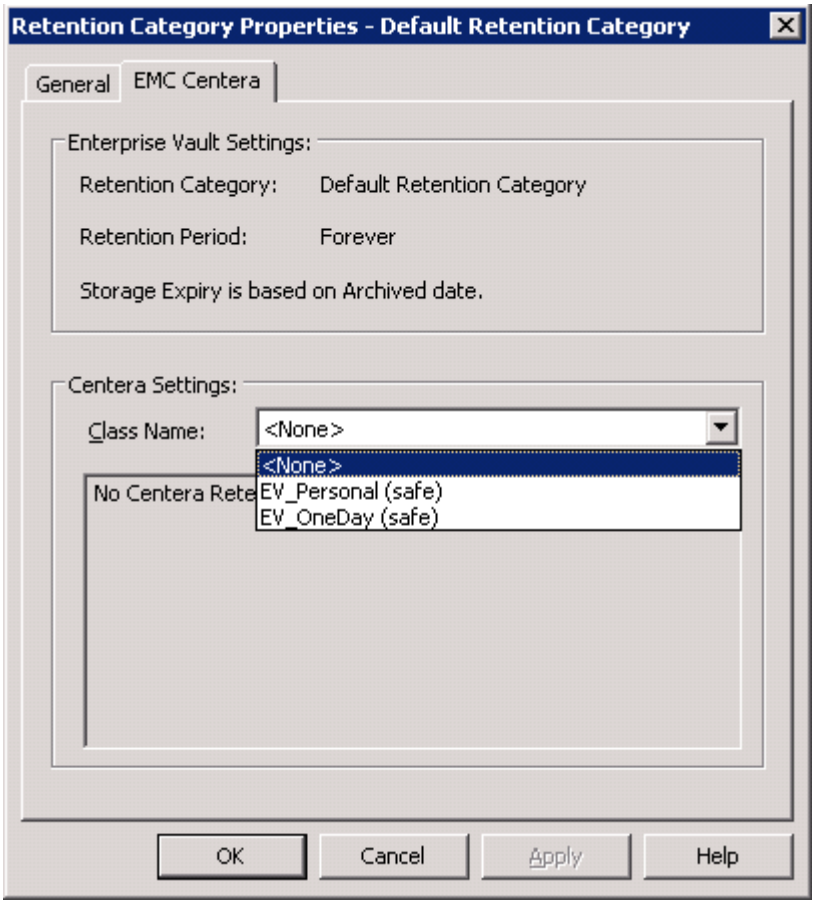

Figure 15 Associating Enterprise Vault Retention Categories and EMC Centera Retention Classes

Policy definition

A policy defines which documents are to be archived and how they are to be archived.

Enterprise Vault creates policies automatically. The Getting Started wizard uses the default Enterprise Vault policies. The default policies in Express mode and Custom mode are the same and, if required, can be modified at a later time, using the Administration Console

Custom configuration of Enterprise Vault and EMC Centera best practices

Once the Getting Started wizard has finished, there are still some particular settings pending configuration.

EMC Centera Connection String

EMC Centera Application Authentication is the process whereby an application (in this case Enterprise Vault) has to provide authentication information to EMC Centera before access is granted. The connection string is a parameter that is used by applications when they connect and authenticate to a cluster. In its most basic form, it consists of a number of IP addresses all belonging to the same cluster plus credential information.

Although one IP address will be enough for the SDK to discover all available IP addresses in the primary and replica clusters, it is recommended to provide the IP addresses of all access nodes in the primary cluster to ensure that a connection can be made even if some access nodes are not available (offline). The credential information is contained in the PEA file.

Note: The Enterprise Vault VAC does not require that the IP addresses of the replica Centera are entered; these are always obtained from the primary Centera.

Additional information on Connection strings, authentication, and probing mechanisms is available in the EMC Centera online help, the *EMC Centera Programmer's Guide* and the *EMC Centera SDK Version 3.2 API Reference Guide.*

The PEA file must be accessible to the Enterprise Vault Storage Service that manages the Vault Store Centera partition that will use the profile.

The Storage Service runs under the Vault Service account and therefore the PEA file must allow read access by this account. For greater security access to the PEA file by any other account could be disabled.

The Vault Store Partition's Connection tab allows the user to establish the IP Address (es) for the primary Centera. In addition, a PEA file must be specified; using the Browse button instead of entering the path location will further ensure that the PEA file is accessible (see Figure 16).

Use the Test Settings button to verify the connection to the primary Centera.

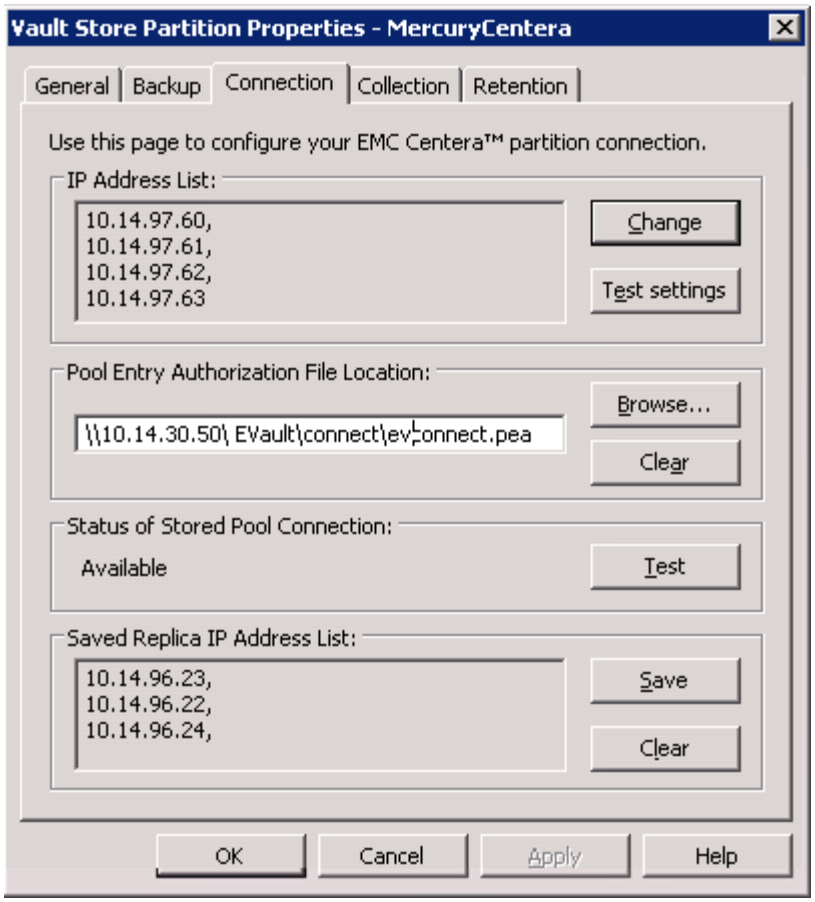

Figure 16 Establishing the EMC Centera connection string

EV version 6.0 supports setting Centera PEA files in the Vault Administration Console. The process is documented in the VAC help as well as in the EMC Centera Global Services Documentation online help.

Network segmentation

In a purely segregated environment of application and replication data traffic, where network segmentation has been established, built-in read failover functionality will not function as expected. For any application that requires read failover capabilities to a secondary EMC Centera, under these circumstances, it is recommended that these failover IP addresses are explicitly specified by adding the access role to the replication nodes in the replica cluster. Reviewing the application settings and ensuring that the application traffic network has access to the secondary EMC Centera is strongly recommended. (For additional information, see the white paper *EMC CentraStar 4.1 Network Segmentation – A Detailed Review*.)

Note: Additional information is available on EMC Centera DIMS 216541.

Threading

A thread is a part of a program that can execute independently of other parts. CentraStar supports multi-threading, thus enabling threads to run at the same time without impacting each other. A single thread cannot take advantage of this parallelism.

In EMC Centera, threads are distributed evenly over the available nodes and the number of nodes influences the number of threads that can be supported.

BEST PRACTICE: It is recommended not to have more than 20 concurrent threads per access node.

Enterprise Vault uses multi-threading to increase the maximum transfer rate by leveraging EMC Centera parallelism capabilities. When archiving with collections, this is not relevant because it is only collections that are written to EMC Centera and not individual items. However, when archiving to EMC Centera without collections, optimum performance is reached when the number of processes is increased. For example:

Number of storage archive processes	Number of PST migrators
10	20

Expanding the Enterprise Vault Servers branch and selecting the Advanced tab, on the Storage Service properties, allows the configuration of the number of threads for writing (used by the Archive processes) and reading (used by the Restore processes) purposes.

When collections are not enabled, the rate at which items are expired and deleted from EMC Centera can be improved by increasing the number of Expiry threads to its maximum value of 10. This is done by changing the value of the following registry key:

Key name: HKEY_LOCAL_MACHINE\SOFTWARE\KVS\Enterprise Vault\Storage

REG_DWORD value: EMCCenteraExpiryThreads

Note: The default value is 5. The maximum number of threads is 10.

BEST PRACTICE: The number of archive processes should be 10 when collections are not used - and the default (5) when collections are used. For Restore it should be set to 1 Restore process and 10 threads, regardless of whether or not collections are used.

The settings displayed in Figure 17 are per Vault Store server.

Note: If there are multiple servers deployed, these settings must be checked for consistency in all of them.

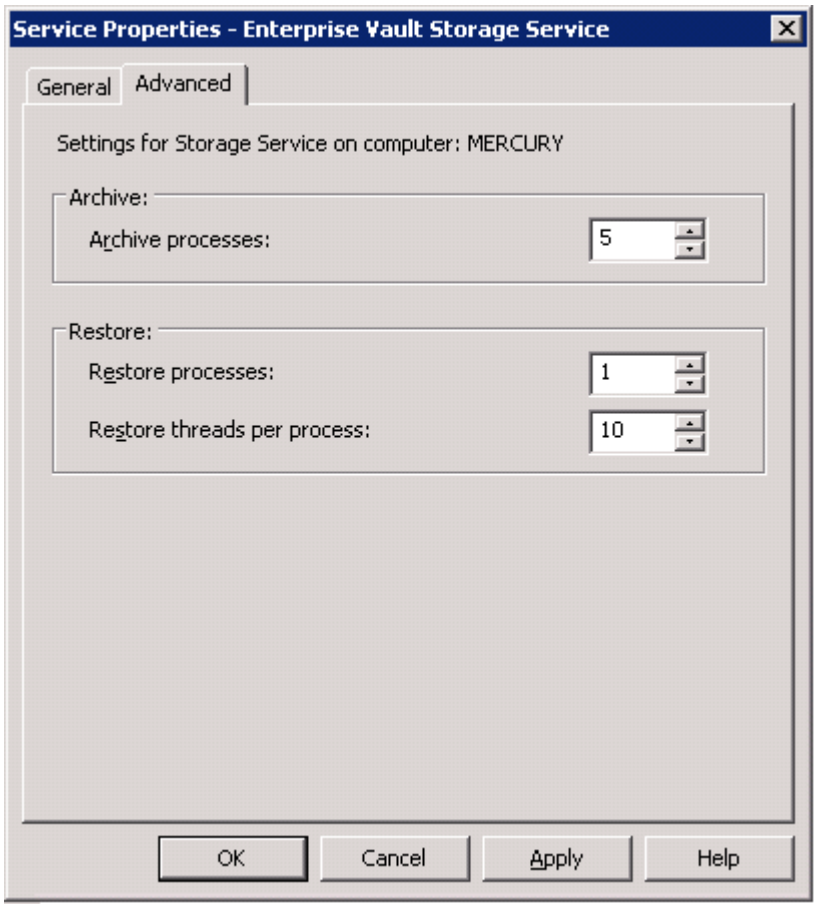

Figure 17 Number of threads

Enterprise Vault retention categories

This graphic shows how to assign a retention category to a provisioning group.

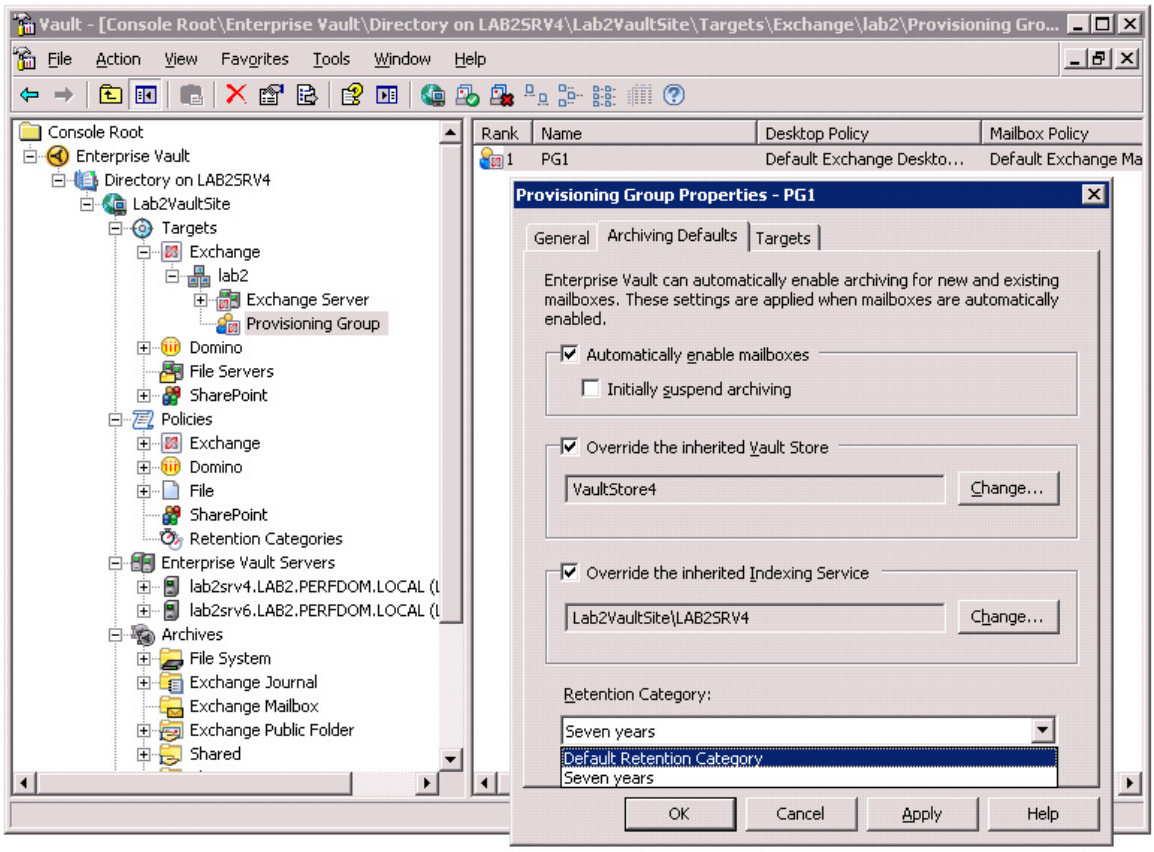

Figure 18 Provisioning Group Properties

Enterprise Vault fine tuning

Additional information for customizations and best practices around mailbox archiving, public folder archiving, File System Archiving and placeholder shortcuts, and performance tuning is available in the *Symantec Enterprise Vault 8.0 - Administrators Guide*

In particular, the following sections are presented in greater detail for mailbox archiving:

- Archiving based on age
- Archiving based on quota or age and quota
- Archiving items from Exchange managed folders
- Archiving items only if they have attachments
- Customizing the Enterprise Vault settings for a journal mailbox
- Disabling archiving for mailboxes

Installation validation

Once Enterprise Vault has been configured, it is recommended to validate that all settings and configurations initially planned have been taken into consideration. The following are the Enterprise Vault and EMC Centera tools and procedures that are available.

Enterprise Vault

Enterprise Vault Deployment Scanner reviews the configuration of a computer and reports on any issues that may stop Enterprise Vault from running on it. In particular, the Test named **EMC Centera Connectivity** checks the connectivity to EMC Centera access nodes.

To run Enterprise Vault Deployment Scanner, the following steps must be performed:

1. Log in using the Vault Service account.
2. Use one of the following methods to start Deployment Scanner:
   - On the Windows Start menu, click **All Programs > Enterprise Vault >Deployment Scanner**.
   - On the Tools menu of the Vault Administration Console, ensure that **Advanced Features** is selected. Then, in the left pane of the Vault Administration Console, right-click the Enterprise Vault server and click **Deployment Scanner**.
3. In the Welcome page, select **Do not set configuration options for Deployment Scanner** if the Deployment Scanner was previously run and you want to rerun it without entering the configuration options again.
4. Click **Next** and then, if **Do not set configuration options for Deployment Scanner** was selected, go to step 6. Otherwise, complete the fields in the two Server Configuration pages.

   Identify the machines on which SQL Server has been installed. Enterprise Vault uses SQL Server to store configuration data and information about the archives.

   If the default database instance is not being used on SQL Server, the instance name must also be entered here, in addition to the name of the database server, in the format sql_server\database_instance. An example is sql\vault. The SQL server(s) identify the Microsoft Exchange servers from which the user wants to archive the items in user mailboxes, journal mailboxes, or public folders.

   - **Microsoft Exchange Server(s)** - Specify the addresses of the Microsoft SharePoint servers that hold the documents to be archived.
   - **Microsoft SharePoint Server(s)** - Identify the file servers that contain files for archiving.
   - **File Server(s)** - Specify the paths to the network shares that contain files for archiving.
   - **File Share(s)** - Specify the IP addresses of access nodes in an EMC Centera cluster.
   - **EMC Centera Access Node(s) -** Specify the IP addresses of access nodes in an EMC Centera cluster.

   In each case, you can add a new item by typing its name or address and then clicking **Add**. To remove an item from a list, click it and then click **Remove**.

5. When you have set all the options, click **Next** to proceed with the tests.
6. Wait a few moments for the utility to perform the tests and display the results.
7. Click the blue links to display more information on each test result. As well as displaying the test results on-screen, Enterprise Vault Deployment Scanner saves the report to an HTML file in the Reports subfolder. The name of the report file identifies the date and time at which you created it.
8. Click **Finish** to exit the Enterprise Vault Deployment Scanner.

EMC Centera

The EMC Technical Solutions Test and Acceptance procedure is performed against the EMC Centera to ensure all configuration parameters are configured and functioning properly. The test and acceptance procedure validates:

- The application can make a successful connection to the virtual pool.
- Retention is set properly by viewing the metadata of a file written by the application.
- EMC and the administrators are receiving daily health reports, alerts, and SNMP traps.
- External network ports on EMC Centera are negotiating correctly.
- Replication is working by testing it with JCenverify.

Chapter 3 Operations

This chapter presents these topics:

EV updates and upgrades 84
EMC Centera upgrades 84
Disaster recovery 86
EMC Centera Backup 88
Monitoring 88
High availability 89

EV updates and upgrades

Day-to-day management tasks

The following provides a checklist of the main day-to-day administration tasks required to maintain optimal performance of the Enterprise Vault system:

- Monitoring Enterprise Vault Services and tasks
- Starting or stopping tasks or services
- Checking logs
- Monitoring Exchange Server journal mailboxes and Domino journal databases
- Monitoring disk usage
- Monitoring MSMQ queues
- Maintaining SQL databases
- Backing up Vault Stores
- Enable archiving for new Microsoft Exchange Server mailboxes or Domino mail files
- Importing PST files (Personal Folder files)
- Importing NSF files
- Changing disks for journal archives
- Monitoring licenses
- Modifying the list of who has access to an archive that is being shared by a number of users

Full details of how to perform these tasks are given in the *Symantec Enterprise Vault 8.0 - Administrators Guide* There are a number of utilities available for performing a variety of tasks, such as re-creating FSA Placeholder shortcuts on a file server, moving archived data from an NTFS device to an EMC Centera device, and managing FSA archive points. For additional information, see the *Symantec Enterprise Vault 8.0 - Utilities Guide*.

EMC Centera upgrades

Customers should not face any compatibility issues when upgrading any of the EMC Centera components of their archiving solution.

EMC strongly recommends upgrading to the latest hardware and software versions available. Based on the EMC Centera SDK backward compatibility with CentraStar, there should not be any issues upgrading CentraStar to a newer version. All ISV applications communicating with an upgraded EMC Centera cluster should still be able to function as expected.

Note: Upgrading to a new CentraStar version does not imply or require customers to upgrade their applications.

EMC Centera partners release their products with the particular EMC Centera SDK version with which the application was developed. Upgrading to a newer ISV application version most likely means upgrading the EMC Centera SDK version it ships with.

Based on the SDK forward- and backward-compatibility statements from EMC, there should not be any EMC Centera SDK/CentraStar compatibility issues when upgrading an application to a newer version.

Note: Upgrading an application version does not imply or require customers to upgrade their Centera CentraStar version. For additional information, see the white paper *EMC Centera CentraStar / SDK Compatibility with Centera ISV Applications*.

In addition to this, as part of their regular business practices, Symantec performs additional testing when major CentraStar releases are available; see the Enterprise Vault Compatibility Chart for detailed information.

Note: Additional information is available in the *Updating EMC Centera Access Node IP – Best Practices Planning* white paper.

Added capacity

Capacity is added to the EMC Centera at the cube level. A cube can consist of up to 16 Gen4 nodes. A cluster can contain up to eight cubes or 128 nodes. Cubes are connected through an interconnection switch. Nodes are added to a cube in two-node expansion increments when using CPM, and eight-node expansion increments when using CPP. The capacity add is automated and is transparent to the application.

CentraStar upgrades

CentraStar upgrades can be performed online. One node is upgraded and rebooted at a time, making an upgrade seamless to the application.

EMC Centera migrations

In EV 6.0 or later multiple EMC Centera partitions can exist in a single Vault Store, which can be useful during EMC Centera hardware migrations.

In this case, using the Vault Administration Console (VAC), all existing partitions pointing to the old EMC Centera will be closed, and new partitions will be created to archive against the new EMC Centera for all on-going data. With this configuration, archiving activities will not be impacted and users will be able to read from both closed and open vaults.

BEST PRACTICE: How to migrate EV data from older to newer Centera hardware is as follows:

1. "Close" all current vaults, so that they are in read-only mode. Consequently, no extra data is added to the old cluster, making this environment stable.

2. Create as many vaults as currently exist, but pointing to the new cluster. All new writes/reads go there; the user should implement the virtual pools, and all other security configurations at this time (pools/profiles/PEA files, and so on).

   Now that the new environment is up and running, it is possible to start the migration process, one vault at a time. This allows the company more control over the process, prioritizing the vaults to migrate first. For each Vault:

3. Obtain the list of clips (savesets) to migrate (by vault). There are two ways for doing this:

   EV support tells the customer the list of clips or the SQL Script to run against EV SQL database

   The customer generates the list using the EVSVR Utility for Enterprise Vault, included in EV 8.0 SP1 or newer.

4. The customer provides a clip list to EMC Centera (Support or Professional Services) to perform the migration.

5. Once migrated and a reconciliation has been completed, the old vault, which was closed, is reconfigured to point to the new EMC Centera (connection string).

6. Repeat steps 3 - 5 as needed.

7. Once all vaults are migrated, the old cluster can be disposed.

Disaster recovery

EMC Centera Replication provides a disaster recovery mechanism for content written to EMC Centera clusters. EMC Centera Replication can be used to create multiple protected copies of content written to a primary Centera cluster by automatically copying the content to a replica Centera cluster. Replication runs as an asynchronous background tasks and can be configured in a number of topologies, namely unidirectional, bidirectional, chain, star, and star and chain. (For detailed information on disaster recovery, refer to the *Symantec Enterprise Vault 8.0 - Administrators Guide*

EMC Centera Replication provides functionality that allows the recovery of the content that is either missing or simply unavailable. Applications will fail over automatically to retrieve missing or unavailable data content on the primary cluster from the replica cluster.

Additional information for EMC Centera Replication is available in the *EMC Centera Replication - A Detailed Review* white paper.

The following sections describe the DR procedures for events impacting either Enterprise Vault or EMC Centera. If the event impacted both components of the environment, follow both procedures.

Enterprise Vault only disaster recovery

If a disaster event impacted only the Enterprise Vault portion of the archival infrastructure, following the methodology and procedures described in the “Recovery” section of the *Symantec Enterprise Vault 8.0 - Administrators Guide* as it applies to the strategy chosen (for example, full system backup, data-only backup) should get the application up and running.

Regardless of the strategy used, it is necessary to ensure that the Pool Entry Authorization (PEA) file(s) is part of the backup content checklist plan, as it grants access to the EMC Centera cluster (for additional information see Centera Connection).

The following is the list of activities that needed to perform to re-establish normal operations:

1. For a full system backup strategy, the user may want to Test the connection from the Vault Partition Properties Connection tab (Figure 16).

2. In the case that a full system backup strategy is not used, while following the recovery procedures described in the *Symantec Enterprise Vault 8.0 - Administrators Guide* restore the PEA file and provide its new location path at the time of executing the Enterprise Vault Configuration, and follow the best practices laid out in this document.

3. Due to changes made since the last set of backups were done and because certain operations may not have completed before the system failure occurred, revert all pending shortcuts in order to re-archive them (See Step 6 “Repeat Operations” in the “Recovery” section of the *Symantec Enterprise Vault 8.0 - Administrators Guide* for more details):

   – Repeat archive operations done since the last set of daily backups were made.

   – Cancel all archive pending items from mailboxes. Symantec’s tech note 273175 describes "*How to cancel items in an "archive pending" state in Enterprise Vault (EV) for Microsoft Exchange*"; Solution 3 - using the VAC to change the Pending Shortcut Timeout to 0 – is the recommended method.

4. Optional: To further verify the integrity between the Enterprise Vault database(s) and the EMC Centera archive, use the EVSVR Utility provided with the Enterprise Vault 8.0 Installation kit. Additional information on the use of the tool is available in the *Symantec Enterprise Vault 8.0 - Utilities Guide.*

Note: Enterprise Vault 8.0 SP2 is the first version that includes repair operations as part of EVSVR. Corrective actions are available to Symantec Support in previous versions of Enterprise Vault, via platform specific tools; EVCentera Checker could be used by Symantec support if inconsistencies are found.

EMC Centera primary only disaster recovery

If a disaster event impacted only the Centera primary cluster, Enterprise Vault will automatically fail over for reads.

Due to the asynchronous nature of the EMC Centera Replication mechanism, any data written to the primary Centera still resides on primary storage (Exchange Server) as part of the Safety Copy functionality, until its replication has been confirmed. Since the data written to the primary EMC Centera might not have been replicated prior to the occurrence of the disaster event, it is necessary to re-archive any data still identified as “Shortcut Pending” by the Safety Copy, in order to sync up the system.

To allow full recovery, it is necessary to reconfigure the EV server(s) to point to the new (former replica) Centera as the primary cluster, and then re-archive the data:

1. Look at the Connection tab properties of the Vault Store and add the “Saved Replica IP address List” as the “IP Address List” (Figure 16).

2. Because certain operations may not have completed before the system failure occurred, revert all pending shortcuts in order to rearchive them (See Step 6 “Repeat Operations” in the “Recovery” section of the *Symantec Enterprise Vault 8.0 - Administrators Guide* for more details):

   – Repeat archive operations done since the last set of daily backups were made.

   – Cancel all archive pending items from mailboxes. Symantec’s tech note 273175 describes "How to cancel items in an "archive pending" state in Enterprise Vault (EV) for Microsoft Exchange"; Solution 3 - using the VAC to change the Pending Shortcut Timeout to 0 – is the recommended method.

3. Optional: To further verify the integrity between the Enterprise Vault database(s) and the Centera archive, use the EVSVR Utility provided with the Enterprise Vault 8.0 Installation kit. Additional information on the use of the tool is available in the *Symantec Enterprise Vault 8.0 - Utilities Guide*.

Note: Enterprise Vault 8.0 SP2 is the first version that includes repair operations as part of EVSVR. Corrective actions are available to Symantec Support in previous versions of Enterprise Vault, via platform specific tools; EV CenteraChecker could be used by Symantec support if inconsistencies are found.

EMC Centera Backup

Tape out (Seven10 Storfirst Altus)

- Altus is a complete tape library and VTL management solution and manages tape and VTL data in all possible locations: in the drive, in the library slot, and offline (on-shelf or in-vault).
- Altus is designed specifically to back up an EMC Centera to tape or a virtual tape library.
- Altus builds a single, complete, and always-synchronized copy of all the information in EMC Centera by incrementally capturing new C-Clips
- Altus can perform a full, partial, or a single clip restore.

Monitoring

Centera Console

Centera Console is a Web-based user interface that enables system administrators to view detailed information concerning the health, capacity, and performance of one or more clusters in their EMC Centera environment

CV/CLI

CV/CLI is a Java-based tool used to monitor and manage an EMC Centera.

Health report

A daily HTML report is e-mailed to system administrators reporting the state of the EMC Centera.

Alerts

An XML report is e-mailed to system administrators reporting on an alert that had occurred on the EMC Centera.

Audit logging

EMC Centera automatically logs information that allows the system administrator to see who logged in to the EMC Centera and which actions were performed.

High availability

For detailed information on configuring a working building blocks solution and recovery, refer to the *Symantec Enterprise Vault 8.0 - Administrators Guide*

Chapter 4 Performance Expectations

This chapter presents these topics:

Enterprise Vault service level agreement (SLA) 92
Performance benchmarks 92
Cooperative Support 95

The performance of Enterprise Vault with EMC Centera has been extensively tested both at Symantec performance test laboratories and at EMC performance test laboratories.

The conclusion of the testing is that EMC Centera is very scalable and well suited to large enterprises.

Enterprise Vault service level agreement (SLA)

Table 11 lists Symantec's expected throughput for different numbers of physical CPUs per server, having an average message size of 70 KB, including attachments.

Table 11 Expected Enterprise Vault ingest rates for physical CPUs

Number of CPUs	Hourly ingest rate (70 KB)
2	25,000
4 single or 2 dual processors	40,000
2 quad-core processors	60,000

Note: The average size of mail messages has an effect on the throughput. The observed effect is that when the average message size is doubled, throughput is reduced by one-third.

Performance benchmarks

A series of joint performance and stress tests have been conducted by Symantec and EMC.

Table 12 shows the results from the tests of March 2007. Dual-processor computers hosted the Enterprise Vault servers.

Table 12 Performance and stress test results

Enterprise Vault task	Enterprise Vault SLA	EMC Centera benchmark	Delta
Normal Tests			
E-mail archiving (msg/hour)	40,000	50,000	25%
Journal archiving (msg/hour)	40,000	40,000	0%
Journal and archive (msg/hour)	N/A	48,000	N/A
File System Archive	40,000	47,000	17%
PST migrations	40,000	56,000	40%
Online viewing	40,000	100,000	150%
PST exports	30,000	50,000	60%
Storage Expire	100,000	160,000	60%
Re-indexing	100,000	100,000	0%
Stress Tests			

Enterprise Vault Storage API	N/A	1,000,000	N/A

In this test, a business use case to be able to archive and back up the content from 50,000 mailboxes, each of them having 20 messages eligible for archive was proposed; this translated to 1 million messages to be archived on a daily archiving window timeframe.

The testing proved that enterprises using Enterprise Vault and EMC Centera as a joint e-mail archiving solution will be able to comfortably archive 50,000 user mailboxes in less than three hours; further savings on the operational overhead required for unnecessary backup of the archive was reinforced, as replication was implemented. In the case of extraordinary events that force the execution of both journal archiving and e-mail archiving simultaneously, customers will be able to do so in less than 4 hours.

At the same time, it was proven that archiving to EMC Centera meets or exceeds Enterprise Vault service level agreements and provides a resilient infrastructure against failures in either component of the solution.

Infrastructure components for this test that refer to the building block are listed in Table 13.

Note: The best practices mentioned in this paper were followed.

Table 13 Performance test components configuration

Component (number of computers)	Configuration
SQL Server (1)	SQL Server 2005 on quad Intel Xeon (2.7 GHz), 8 GB RAM
Exchange Server (8)	Exchange Server 2003 on dual Intel Xeon (3.4 GHz), 4 GB RAM
Symantec Enterprise Vault (8)	Exchange Server 2003 on dual Intel Xeon (3.4 GHz), 4 GB RAM
Primary storage array (replicated)	EMC CLARiiON® CX3-40
EMC Centera archive (replicated)	Centera Gen4, 16 nodes. Roles: 4 dual Access/Storage, 12 Storage

Ingest

Archival activities involve journal e-mail archiving and regular e-mail archiving

- Seven Enterprise Vault servers were able to archive at a rate of about 350,000 items an hour into the EMC Centera (or 50,000 items per server). All items had been replicated and indexed within a short time of the end of the test.
- Eight Enterprise Vault servers were able to journal at a rate of about 320,000 items an hour (or 40,000 items per server).
- Eight Enterprise Vault servers were able to simultaneously journal and archive at a rate of about 390,000 items an hour (or about 48,000 items per server).

Note: Although rare, two ingest methods running at the same time might be required after a regularly scheduled archiving activity task has been postponed or aborted.

- Eight Enterprise Vault servers were able to archive files at a rate of about 380,000 items an hour (or about 47,000 items per server).
- Eight Enterprise Vault servers were able to migrate (archive) PSTs at a rate of about 450,000 items an hour (or about 56,000 items per server). During the migration, the CPU usage of the EMC Centera access nodes was 50% and in the Centera storage nodes it was 33%.

Note: EMC Centera could have sustained a much higher throughput but it is not possible to extrapolate an accurate maximum number.

Note: All forms of ingest generated from Enterprise Vault access the EMC Centera in the same way. A rate of 450,000 items an hour represents the maximum measured ingest rate for all forms of ingest tested. The ingest rate was limited by the number of Enterprise Vault servers available for testing. The absolute maximum is much higher than this, but it is not possible to speculate what this may be.

Retrieval

Retrieval tests were set up so that each retrieval request was for a different item avoiding any effects of caching either in EMC Centera or Enterprise Vault.

- Eight Enterprise Vault servers were able to download items at a rate of about 100,000 items an hour. The Enterprise Vault servers were simulating online viewing user requests to download data – for example when clicking on a shortcut. The average server response time was about 360 milliseconds; this is the time required for Enterprise Vault to download the item together with any attachments and prepare the item for viewing. For the user, the response time will be this time plus the time to transfer and render the item on the user's desktop.
- Eight Enterprise Vault servers were able to export items to PST files at a rate of about 400,000 items an hour (or about 50,000 items per server).

Storage Expiry

Eventually, retention periods expire and items become eligible for deletion.

- Eight Enterprise Vault servers were able to expire (delete) items at rate of about 1,300,000 items an hour (or about 162,000 items per server).

Re-indexing

- Under some circumstances, AltaVista indexes need to be rebuilt. This means that EMC Centera needs to be accessed to gather the necessary indexing information.
- A single journal index re-indexed at a rate of 100,000 items an hour while the system was idle and 60,000 items an hour when journaling was active at the same time. The re-indexing had no effect on the journaling rates. The slowdown in the re-index rate was due to contention on the EV server and not on the EMC Centera.

Accelerator applications

EMC Centera is an ideal storage medium when using the Enterprise Vault Accelerator products. The following conclusions were reached from the tests:

- For up to 100 concurrent reviewers (each selecting a new message every 30 seconds), the response time, which includes downloading the item from EMC Centera, was less than 0.5 seconds. Over 100 users, response times began to get longer but this was related to resource and configuration issues on the Enterprise Vault servers and not EMC Centera.

Note: Many reviewer actions do not download items and these are unaffected by the storage medium.

Stress tests

All retrieval request access EMC Centera in the same way, for example, user initiated downloads, re-indexing, export archive, and retrieving items for review by the Accelerator applications. In an attempt to tax EMC Centera, the Enterprise Vault Storage API was isolated from the application context, and run directly against EMC Centera.

- Eight servers, using the Enterprise Vault Storage API, were able to retrieve about 1,000,000 items per hour.

Note: This represents the maximum retrieval rate for all retrieval methods given enough Enterprise Vault servers. During the retrieval the CPU usage of the EMC Centera access nodes was close to 100% and the CPU usage of the EMC Centera storage nodes was 30%.

Cooperative Support

EMC and Symantec manage its collaborative support in two different ways, through Cooperative Support Agreements (CSAs) and through TSANet, a web-based customer support consortium. EMC is a member of TSANet, and holds a seat on its board of directors, along with more than 100 other technology companies. Symantec and EMC use TSANet to collaborate on interoperability issues. As part of TSANet, companies are members of a Mission Critical Community that enables its members to triage and collaborate on issues on a 24/7 basis until a call owner for problem resolution is determined. EMC and Symantec, in conjunction with Oracle, have also created a Multi-Vendor Escalation Center (MVEC) for the purposes of providing seamless backline technical support. The center provides another level of collaboration between the support organizations that ultimately results in a better customer experience. (For additional information, see the *Symantec Enterprise Vault 8.0 Performance Guide*.

Chapter 5 Support Model

This chapter presents these topics:

Symantec Enterprise Vault considerations 98
EMC Centera considerations 99

Symantec Enterprise Vault considerations

If you are an existing Enterprise Vault user who is experiencing problems because of environmental factors, use Deployment Scanner to collect and export configuration information in a form that Symantec Technical Support can analyze.

To export information about the environment:

1. On the Welcome page of Deployment Scanner, select **Gather information about the environment**, and then click **Next**.
2. Select the environment checks to perform, and then click **Next**.
3. On the Support Case page, enter the support number if one exists, and choose where to save the exported data.
4. Click **Next** to collect and export the information about the environment.

Enterprise Vault logging with DTRACE

DTRACE is a command-line utility that logs what an Enterprise Vault Service, process, or task is doing at the code level and provides a way to diagnose what is going wrong. With DTrace, it is possible to monitor multiple services simultaneously, filtering for specific words and writing the trace to a log file.

The use of this tool is commonly one of the first steps required when logging a support call with the Symantec Enterprise Vault support team.

To use DTRACE to diagnose an EMC Centera-specific problem, it is necessary to run DTRACE manually from Start > Programs > Enterprise Vault > Dtrace.

Typing **View** from the command line will display the components that can be traced; EMC Centera calls may be made from several of these components. Depending on the nature of the problem, the following may need to be traced:

- ArchiveTask - Exchange archiving
- EvLotusDominoArchivingTask - Domino archiving
- EvSharePointArchiveTask - SharePoint archiving
- EvFsaArchivingTask - FSA archiving
- StorageArchive - Any archiving without collections
- StorageFileWatch - Any archiving with collections
- StorageRetrieval - Retrieval of items
- StorageOnlineOpns - Retrieval of items by end user “on-demand” requests
- StorageCrawler - Retrieval of items for indexing
- StorageDelete - Expiry of items, deletion of archives and Vault Stores, and NTFS to EMC Centera migrations

Note: Additional information on the processes involved when Centera Collections are enabled is available in the "Collections" section of the document.

To set the components to be traced, the following commands are recommended:

- **Set nn verbose** - Where *nn* is the component id returned from the View command.

To set the trace to log only EMC Centera commands:

- Filter clear both
- Filter set Centera

To start monitoring and direct output to a log file:

- **Log Centera.log** - Will create a log file in the Enterprise Vault program folder called Centera.log and will start logging.

While DTRACE is running, the components that are enabled will run slower. The log file may grow rapidly. Normally DTRACE should not run for more than 10 minutes (but may need to run longer to catch a more infrequent error).

The log file may be mailed to the support team handling the problem.

EMC Centera considerations

SDK logging

SDK logging is the process of capturing SDK activity between the application and EMC Centera. This log is used for SDK troubleshooting and development. For application debugging of supported platforms, the SDK provides thread-safe logging of all its activities via log API functions or logging environment variables. Logging does not create new threads.

Note: When using environment variables, an application server restart is required before and after the log file is collected. Enterprise Vault does not use API functions and relies solely on the environment variables mechanism to enable SDK logging.

When checking for log state settings, the FPLogging mechanism observes the following order of priority, from highest to lowest:

1. dfd FPLogState settings applied by FPLogging_Start() calls
2. FPLogState.cfg in the working directory
3. FP_LOG_STATE_PATH environment variable
4. Logging environment variables

For example, during application startup, FPLogging first checks for the presence of the FPLogState.cfg properties file in the working directory. If this file exists, the SDK automatically reads and uses the settings contained in that file, including the log path for the logging output. No attempt is made to read any logging environment variables unless FPLogState.cfg is absent from the working directory.

If this is the case and environment variables are to be used, the log state settings of the configuration file defined in FP_LOG_STATE_PATH (if specified) take precedence. That is, the SDK ignores the log path specified in the FP_LOGPATH environment variable.

Note: To avoid any confusion, it is recommended to verify that the FPLogState.cfg file is not present in the working directory. Table 14 lists the required environment variables.

Table 14 Required SDK environment variables

Environment variable	Recommended value	Description
FP_LOGPATH	Example: C:\mylog.txt	The full path and file name of the file to receive the log data and restart the Revo app

The following are the steps to enable and collect SDK logging using enterprise variables:

1. Create the FP_LOGPATH environment variable with the valid path

Note: Log files can grow quite rapidly. It is recommended to disable logging as soon as the issue has been reproduced

2. Restart Enterprise Vault server
3. Reproduce the issue
4. Remove the environment variable
5. Restart the application

Once the SDK log file is collected, it must be sent to EMC Customer Support for analysis.

Additional information on available logging mechanisms and environment variables is available in the *EMC Centera SDK Version 3.2 API Reference Guide*.

Application Registration

Application Registration is an automated collection of application information. This information provides instant data for all applications connected to the EMC Centera. Information includes application make, version, hostname, hardware version, operating system, profile used to connect, SDK version, date of first connect, date of most recent connection, and number of successful connections OS.

Audit logging (syslog)

Audit logging provides audit trail security information for all management actions. These actions include:

- Changes made with Centera Viewer (CV), Centera CLI, Management API (MaPI) and Centera Platform commands
- Failed and successful management logins
- Failed and successful SDK logins
- All management actions that result in a configuration change

Audit logs are stored safely in a system pool on the EMC Centera called AuditArchive. Each log entry can be sent out through the syslog industry-standard logging protocol through UDP port 514. The administrator can define how long this log information is kept before it is automatically purged. A human readable audit trail is available via the CV/CLI and via a syslog interface.

Alerts

EMC Centera actively monitors its health, capacity, and performance using sensors. Each sensor has a value that it periodically records by monitoring EMC Centera hardware and software components such as nodes, disks, replication queues, and so on. Fixed rules and thresholds are defined that decide if and when an alert has to be generated. An alert is a message with information on the cluster's state to notify the system administrator and EMC using SMTP and SNMP of a potential problem.

Chapter 6 Conclusion

Maintaining its factory-defaults, EMC Centera provides a quick and simple storage configuration requiring only a custom pool/profile setup to exactly match Enterprise Vault access, and security requirements. This, combined with Enterprise Vault's convenient wizard-like graphical user interface to easily fine-tune the behavior and performance of the archive policy engine when archiving to EMC Centera, makes the joint Symantec/EMC offering an almost "turn-key" alternative for e-mail and other unstructured data type archiving.

Given the robustness, flexibility, and maturity of the solution, it is not the installation and configuration of the Enterprise Vault and EMC Centera that is most critical to the successful implementation of archiving initiatives in the enterprise, but rather the careful planning and design of the architecture. This initial phase must be driven by a clear understanding of the short-, mid-, and long-term stakeholder's needs, wants, and expectations, along with environmental and organizational constraints.

Together, Enterprise Vault and Centera can comfortably scale to satisfy the most demanding and challenging of the environments, as benchmarks and existing deployments alike have proven over time. At the same time, the solution provides a secure yet simple framework for companies to achieve increased operational efficiencies by dramatically reducing the size of e-mail stores and optimizing storage utilization in the face of data growth and longer retention periods. Furthermore, the offering enables companies to obtain risk mitigation associated with growing regulatory and corporate governance requirements for data permanence, security, and confidentiality, and to maintain assured authenticity and easy accessibility to archived records.

Appendix A Glossary

This appendix presents this topic:

Terminology 106

Terminology

Term	Definition
Access Profile	Access profiles are used by applications and users of management tools to authenticate to a cluster, and by clusters to authenticate to another cluster for replication or restore connections. System administrators can create access profiles using the CLI. Each access profile consists of a profile name, a secret (password), and a set of capabilities and roles.
BLOB	A BLOB is a series of bytes that represents a fixed content object stored in EMC Centera. The format and structure of the BLOB are wholly owned by the client application, and neither the SDK nor the cluster attempts to interpret the binary object.
Building block	Represents the required amount of resources required to support a specific number of Exchange 2007 users on a single VM. The amount of required resources is derived from a specific user profile type.
C-Clip (Clip)	A C-Clip is an application object that represents a bundle of fixed content data and metadata.
Call detail record (CDR)	CDR is the computer record produced by a telephone exchange containing details of a call that passed through it. It is the automated equivalent of the paper toll tickets that were written and timed by operators for long-distance calls in a manual telephone exchange.
C-Clip Descriptor File (CDF)	The C-Clip Descriptor File (CDF) is the physical object stored to the cluster that represents an application-defined C-Clip. The CDF contains all of the metadata specified by the client application, and the links to all associated BLOBs. The CDF-to-BLOB relationship can have a cardinality of one-to-one, one-to-many, or many-to-many.
Centera API / Centera SDK	The EMC Centera SDK is a set of cross-platform application programming interfaces (API) that make it simple for customer applications to perform functions such as store, retrieve, delete, and query for data objects in a variety of flexible and powerful ways. All applications must use this API to read and write to EMC Centera.
Centera Capabilities	Pool-bound content access rights granted by the system administrator to an access profile. They determine which operations an application can perform on the pool data. Possible capabilities are write (w), read (r), delete (d), exist (e), privileged delete (D), query (q), clip copy (c), Purge (p), and Litigation hold (h).
Centera CLI	The EMC Centera Command Line Interface (CLI) is a tool for system administrators to manage and monitor EMC Centera.
Centera Cluster	A cluster is a single logical CAS archive that is accessible to an SDK-based client application. Client applications can store, retrieve, and delete fixed content objects from a cluster. A single cluster can be accessed by one or more applications via a set of node IP addresses and access profiles. Clustered nodes are automatically aware of nodes that attach to and detach from the cluster.
Centera Independent Software Vendor (ISV)	Independent software vendor (ISV) is a business term for companies specializing in making or selling software, usually for vertical markets, such as medical imaging or e-mail archiving. An ISV makes and sells software products that run on one or more computer hardware or operating system platforms. In this case, an EMC Centera ISV is a software company that integrates its archiving solution to EMC Centera.
Centera Partner/ISV	An EMC Centera partner is an independent software vendor (ISV) that integrates to EMC Centera. These two terms are used interchangeably throughout this paper.
Centera Software Development Kit (SDK)	The EMC Centera SDK is a set of cross-platform application programming interfaces (API) that make it simple for customer applications to perform functions such as store, retrieve, delete, and query for data objects in a variety of flexible and powerful ways. All applications must use this API to read and write to EMC Centera.
CentraStar	EMC firmware used by EMC Centera.
Cluster Mask	Defines the EMC Centera capabilities that access profiles can enable. At the cluster level, the cluster (authorization) mask is used to override other profiles

Content Address	A data object's unique identifier. A Content Address is the claim ticket that is returned to the client application when an object is stored to the archive.
File System archiving (FSA)	FSA manages rapidly growing data such as office documents, web pages, images, and audio and video files by automatically migrating inactive data to more cost-effective, self-managed archive storage. The result is improved productivity and compliance, plus faster, more reliable recovery.
Metadata	Data about data. Metadata is information about an informational resource, be that a document (such as a webpage), image, dataset, or other resource. Metadata is valuable in the storage and retrieval of information. Resources supported by good quality, structured metadata are more easily discoverable.
MS Exchange Solution Reviewed Program (ESRP)	The ESRP - Storage program was developed by Microsoft Corporation to provide a common storage testing framework for storage vendors to provide information on their storage solutions for Microsoft Exchange Server.
Node	Logically, a network entity that is uniquely identified through a system ID, IP address, and port. Physically, a node is a computer system that is part of the EMC Centera cluster.
Node Role	The roles that can be assigned to each individual node are either external or internal. Nodes with an external node role have an external IP address configured and use their Eth2 port for communication with the customer's network; external roles are access¸ management, and replication. Storage role is the only internal role. Refer to the online help for additional information.
Pool Mask	Defines the EMC Centera capabilities granted to a particular virtual pool.
Production Archive	Production Archives is a term that customers have dubbed for the use of an archive as an extension of primary storage.
Recovery Point Objective (RPO)	RPO describes the acceptable amount of data loss measured in time.
Recovery Time Objective (RTO)	RTO is the duration of time and a service level within which a business process must be restored after a disaster (or disruption) in order to avoid unacceptable consequences associated with a break in business continuity.
SDK backward compatibility	In technology, especially computing (irrespective of platform), a product is said to be backward-compatible when it is able to take the place of an older product, by interoperating with products that were designed for the older product versions. In this particular case, it refers to the compatibility between older SDK versions and newer CentraStar versions.
SDK binary compatibility	The use of a stable application binary interface (ABI) – the low-level interface between an application program and its libraries, or that between component parts of an application – that enables partners to dynamically replace older SDK versions with newer ones without having to recompile their application.
SDK forward compatibility	Forward compatibility is the ability of a system to gracefully accept input intended for later versions of itself. In this particular case, it refers to the compatibility between newer SDK versions and older CentraStar versions. **NOTE:** The *EMC Centera CentraStar and SDK Release and Interoperability Matrix* has additional information.
TCO	Total cost of ownership (TCO) is a financial estimate. Its purpose is to help enterprise managers determine direct and indirect costs of a product or system.
WORM	WORM (write once, read many) is a data storage technology that allows information to be written to a disc a single time and prevents the drive from erasing the data. The discs are intentionally not rewritable, because they are especially intended to store data that the user does not want to erase accidentally. Because of this feature, WORM devices have long been used for the archival purposes.

6.