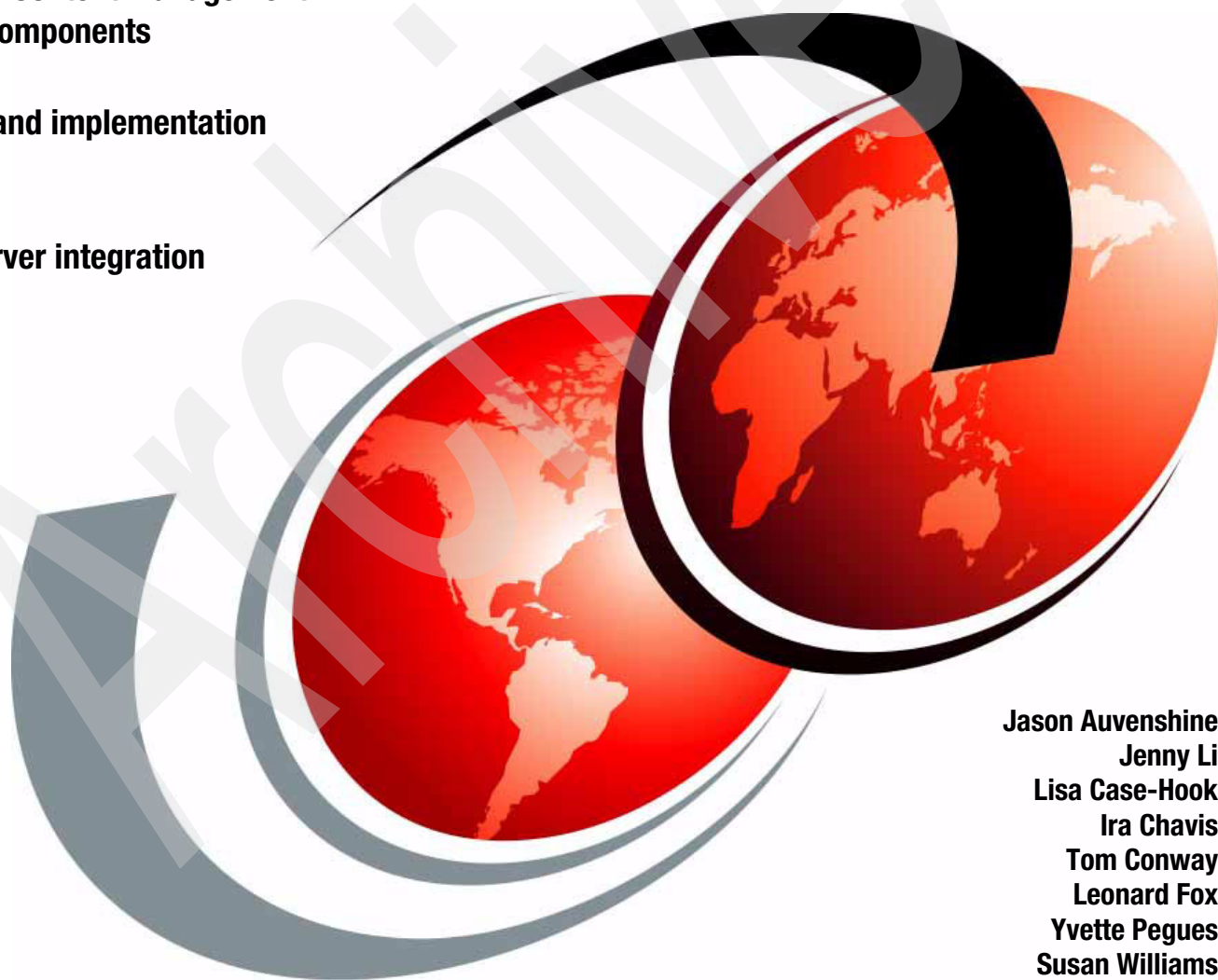**IBM**

IBM Enterprise Content Management and DR550 for E-mail Archiving and Records Management Overview

Enterprise Content Management solution components

Planning and implementation

E-mail server integration

Jason Auvenshine
Jenny Li
Lisa Case-Hook
Ira Chavis
Tom Conway
Leonard Fox
Yvette Pegues
Susan Williams

Redpaper

ibm.com/redbooks

International Technical Support Organization

**IBM Enterprise Content Management and DR550 for
E-mail Archiving and Records Management Overview**

February 2009

Note: Before using this information and the product it supports, read the information in "Notices" on page v.

Second Edition (February 2009)

This edition applies to IBM DB2 Content Manager Enterprise Edition, Version 8.3 and related IBM software.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

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

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

v

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX 5L™	IBM®	Redbooks (logo) ®
AIX®	Lotus Notes®	System p®
BladeCenter®	Lotus®	System Storage™
DB2®	Notes®	System x®
Domino®	PartnerWorld®	Tivoli®
DS4000®	Redbooks®	WebSphere®

The following terms are trademarks of other companies:

SAP, and SAP logos are trademarks or registered trademarks of SAP AG in Germany and in several other countries.

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

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

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

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

Preface

This IBM® Redpaper publication provides an installation and configuration guide for an e-mail archiving reference implementation. This solution leverages the use of DR550 and major IBM software components, such as Content Manager, CommonStore, and Records Manager. The objective is to create and validate the architecture that enables the use of DR550 to support an end to end Information Lifecycle Management (ILM) scenario by demonstrating, with Lotus® Domino® or Microsoft® Exchange, e-mail archiving and retention scenarios.

> **Note:** The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ from the current versions.

This document is intended for IT professionals who are responsible for building and implementing the environment associated with ILM deployments.

The team that wrote this paper

This paper was produced by a team of specialists from around the world working at the International Technical Support Organization, Poughkeepsie Center.

Jason Auvenshine is a Storage Development Architect from Tucson, Arizona. Jason is currently responsible for the development architecture of IBM DR550 and Grid Access Manager products. Jason has over fifteen years of experience working in diverse roles in the Information Technology field. He has worked in Programming, Testing, System Administration, Autonomic Systems, Web Deployment, IT Architecture, and Development Architecture. Jason is an active inventor, and has so far been issued five patents by the US Patent and Trademark Office.

Jenny Li is a Certified IT Architect with the Center for Solution Integration (CSI) team in Poughkeepsie, NY. She is the chair of the IBM Solution Community, which consists of members from solution centers across the company. Her areas of expertise include application architecture for on demand e-business, enterprise solution research, and enterprise architecture and strategy. She is also an active member of the IEEE Computer Society. She can be contacted at mailto:jennyli@us.ibm.com

Lisa Case-Hook is an IT Specialist with the IBM Systems and Technology Group, Center for Solutions Integration (CSI) team in Poughkeepsie, NY, where she performs solution proof of concept and platform architecture validation activities, including the creation of reusable assets for field practitioners. She has over 23 years of experience in IBM, including development, test, and project management aspects of IBM's large and mid-range operating systems. Prior to her recent move to the CSI team, she specialized in performance evaluation and failover testing of complex, leading edge, multi-platform customer systems. She can be contacted at mailto:lisacase@us.ibm.com

Ira Chavis is a Certified Consulting IT Specialist in the Industry Solutions and Proof of Concept Centers in IBM Systems and Technology Group. Working in the Center for Solution Integration, he currently specializes in infrastructure architecture and solutions involving IBM server and storage technologies. Ira has more than 26 years of diversified software

engineering and IT experience. Prior to joining IBM, he worked at Digital Equipment Corporation in various assignments. Ira is an IBM eServer™ Certified Expert in xSeries® and in Grid Technical Sales, a Microsoft® Certified System Engineer (NT4), and a Red Hat Certified Technician. He can be contacted at mailto:ichavis@us.ibm.com

Tom Conway is an Infrastructure Architect with 18 years of experience in the Open Systems Infrastructure field. He joined IBM in 2001 and became the Chief Engineer of the IBM Global Services SAN Interoperability Lab at the IBM National Test Center in Gaithersburg, Maryland. His areas of expertise include Open Systems server hardware, operating systems, networking, and storage hardware and software, including various aspects of Information Lifecycle Management. He is an IBM Certified Professional Server Expert. He can be contacted at mailto:tfconway@us.ibm.com

Leonard Fox is an Advisory IT Architect with the Portals, Content and eCommerce practice of GBS/AIS. He has over 25 years of experience in designing and deploying content management solutions. His experience includes being the lead architect in numerous IBM GBS projects involving Content Management, Records Management, and Document Management. In addition he the technical lead for Document Management in his PCe practice, helping other practitioners configure, propose, and bid on content management solutions using DB2® Content Manager and its add-on components. He can be contacted at mailto:lenfox@us.ibm.com

Yvette Pegues is an Advisory IT Specialist and Storage Consultant with IBM Global Technology Services (GTS). Her experience includes project, program, product, and solution management within the Storage practice and Autobahn (ISD). She is ITIL® Certified and is currently working on the IBM Data Management Services for content management: Content Manager design and Implementation; CommonStore design and implementation; and Records Manager design and implementation for the Storage and Data Management practice. She can be contacted at mailto:yraymon@us.ibm.com

Susan Williams is an IT Specialist with the IBM Systems and Technology Group, Center for Solutions Integration (CSI) team in Poughkeepsie, NY. She has worked for IBM for 14 years and currently designs, develops, and tests solutions for customers. Her job includes the creation of reusable solution assets such as architecture documents, installation and planning manuals, and proofs of concept. Her previous job was in the Electronic Design Automation (EDA) organization, where she performed product design and implementation through system test and customer support. She can be contacted at mailto:susanmwi@us.ibm.com

Thanks to the following for his contribution to this project:

Mike Ebbers

International Technical Support Organization, Poughkeepsie Center

Become a published author

Join us for a two- to six-week residency program! Help write a book dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You will have the opportunity to team with IBM technical professionals, Business Partners, and Clients.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you will develop a network of contacts in IBM development labs, and increase your productivity and marketability.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this paper or other IBM Redbooks® in one of the following ways:

► Use the online **Contact us** review Redbooks form found at:

ibm.com/redbooks

► Send your comments in an e-mail to:

redbooks@us.ibm.com

► Mail your comments to:

IBM Corporation, International Technical Support Organization
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400

Summary of changes

This section describes the technical changes made in this edition of the paper and in previous editions. This edition may also include minor corrections and editorial changes that are not identified.

Summary of Changes
for IBM Enterprise Content Management and DR550 for E-mail Archiving and Records Management Overview
as created or updated on January 21, 2010.

February 2009, Second Edition

This revision reflects the addition, deletion, or modification of new and changed information described below.

Changed information

► The changes in this version represent the transition of the DR550 to a Machine Type Model as well as regular technology refresh.

> **Note:** The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ with the current versions.

1

Executive summary

With business moving at the "speed of now," your users require on demand access to information anytime and anywhere. Businesses run on information, both structured data and unstructured content. Content, which can represent up to 85% of the information an organization processes, includes formats such as documents, images, e-mail, instant messages, digital media, HTML, and XML-based Web content. The challenge is to securely capture, organize, store, and deliver this unstructured information.

1

1.1 Business case

Information management and storage challenges impact an organization's ability to optimize information value and deploy information on demand solutions. The following challenges are inherent in managing and storing information:

- ▶ Surge in criticality, value, and volume of data

    - Projected to grow at an annual rate of 64%.

    - Outpaces IT's ability to collect, store, and manage it by traditional means.

- ▶ Excessive storage costs and missed service level objectives

    - Storage now accounts for more than 15% of total IT budgets.

    - Data growth is now estimated at 25 to 50% annually.

    - In most organizations, effective disk utilization is less than 50%, with 20 to 40% wasted space.

- ▶ Compliance with regulatory requirements and audit procedures

    - There are over 20,000 regulations worldwide that affect data storage, accessibility, and retention requirements

- ▶ Ability to effectively access and gain insight from information once stored, including the growing requirement to support timely economic discovery across existing information

Data that is kept for extended periods of time poses a challenge for many organizations. Today, many companies use non-erasable, non-rewriteable technologies to help protect against accidental or intentional alteration and deletion of information prior to its corporate or regulatory policy-driven expiration. Usually this data is stored on inexpensive archives and repositories, with the trend being toward storage on SATA disk and tape. As organizations address their data retention needs to comply with corporate governance and federal and local regulations (sometimes distilled to record retention schedules), they must keep necessary documents and records available when needed. Overall, the management and control of retention-managed data can be a significant challenge when taking into account other factors, such as costs, integration, performance, security, privacy, and policy-based retention periods.

1.2 Business drivers

These are the major business drivers that lead businesses to adopt an Information Lifecycle Management solution:

- ▶ Reduce costs and simplify management by:

    - Controlling demand for storage

    - Improving asset utilization

    - Reducing hardware, software, and storage personnel costs

    - Reducing data migration effort

- ▶ Enhance services and improve efficiency by:

    - Maximizing and sustaining efficiency by improving the processes and technologies being utilized to deliver storage services to the business

    - Defining and implementing the appropriate storage strategy to address current and future business requirements

    - Enhancing systems and e-mail performance

- – Making better use of existing information
- ► Manage risk and streamline compliance, thereby:
  - – Reducing organizational risk
  - – Complying with governmental regulations
  - – Handling legal discovery or electronic discovery (e-discovery), in which electronic data is sought, located, secured, and searched with the intent of using it as evidence in a legal case

1.3 Solution outline

IBM Data Retention 550 (DR550) is recognized as a leader in the area of data retention management for compliance and governance. It was named "Best in Show" for two years in a row at the AIIM Expo Conference and Exposition in Philadelphia. DR550 is a storage appliance for Information Lifecycle Management (ILM). It bundles an IBM System p® server running AIX®, a DS4000® mid-range disk system, and System Storage™ Archive Manager (SSAM) software. It is ordered and supported as a single Machine Type (2233) with two models (DR1 and DR2). It is delivered to a customer pre-configured in a rack. The DR1 and DR2 models are shown in Figure 1-1.

DR550 Model DR1

DR550 Model DR2

Figure 1-1   DR550 Models DR1 and DR2

The components of a complete e-mail archiving solution include:

► IBM DB2® Content Manager

  Provides the repository for archiving e-mails

► IBM DB2 Records Manager

  Manages the archived e-mail as corporate records

► IBM DB2 CommonStore (Microsoft Exchange or Lotus Domino Notes)

  Integrates with the corporate mail services to extract and archive e-mails from user mailboxes

► IBM System Storage DR550

  A system storage device and media especially configured to archive content

► IBM System Storage Archive Manager

  Archive file management software that integrates the DR550 with Content Manager repository

> **Note:** The baseline configuration of these components is tested and delivered to the client. Thereafter, the client modifies the configuration parameters to meet specific needs beyond the base functionality.

IBM Information Lifecycle Management solution overview

Information Lifecycle Management (ILM) is a combination of processes and technologies that defines when, where, and how data flows through an environment. Adding an ILM solution to a customer's existing environment generates major benefits, including helping users and administrators manage data archiving and retention, and ensuring that data is stored using the appropriate format and media to satisfy availability, security, backup, and compliance requirements as it ages through its life cycle.

The purpose of this publication is to guide the reader towards a Baseline Configuration that validates a recommended set of components required to deploy an IBM Information Lifecycle Management solution for e-mail archiving and retention. Major components used include DR550, Content Manager, CommonStore for Lotus Domino, CommonStore for Microsoft Exchange, and Records Manager. Refer to 2.6, "Reference implementation limitation" on page 17 for the complete listing.

This Baseline Configuration, in itself, does not define a production environment solution for the enterprise. An enterprise deploying an e-mail archiving solution must modify and extend the Baseline Configuration to its own business requirements, and reflecting its own policies and governance needs.

2.1 Assumptions

The Baseline Configuration assumes a distributed multi-platform environment. With the exception of the DR550 Storage System, which is an AIX-based component, the other servers can be deployed on any platform supported by the server component. This includes Windows® OS, UNIX®, Linux®, and AIX. The supported platform "flavors" are documented in the installation guides for each server component.

In this publication, we describe a Microsoft Windows 2003 Standard Server platform for all components except for the Lotus Domino Server running in a Linux platform and the AIX-based DR550 Storage System.

We also assume that the Solution Assurance Review is complete. Refer to the PartnerWorld® Web Site for more details:

http://www.ibm.com/partnerworld/pwhome.nsf/weblook/ben_trs.html

From this page, select the **Technical resources** link to sign in to PartnerWorld. To directly access the Technical resources link, click:

http://www.ibm.com/partnerworld/pwhome.nsf/weblook/ben_detail.html?OpenDocument&type=technical&id=LLIE-6LLS4T

Finally, we assume that all front-end applications (such as Lotus Domino or Microsoft Exchange e-mail server, and Lotus Domino or Microsoft Outlook® mail client) have already been installed in the existing environment.

2.2 Architectural overview

Figure 2-1 on page 7 illustrates the high-level architecture of the solution as implemented.
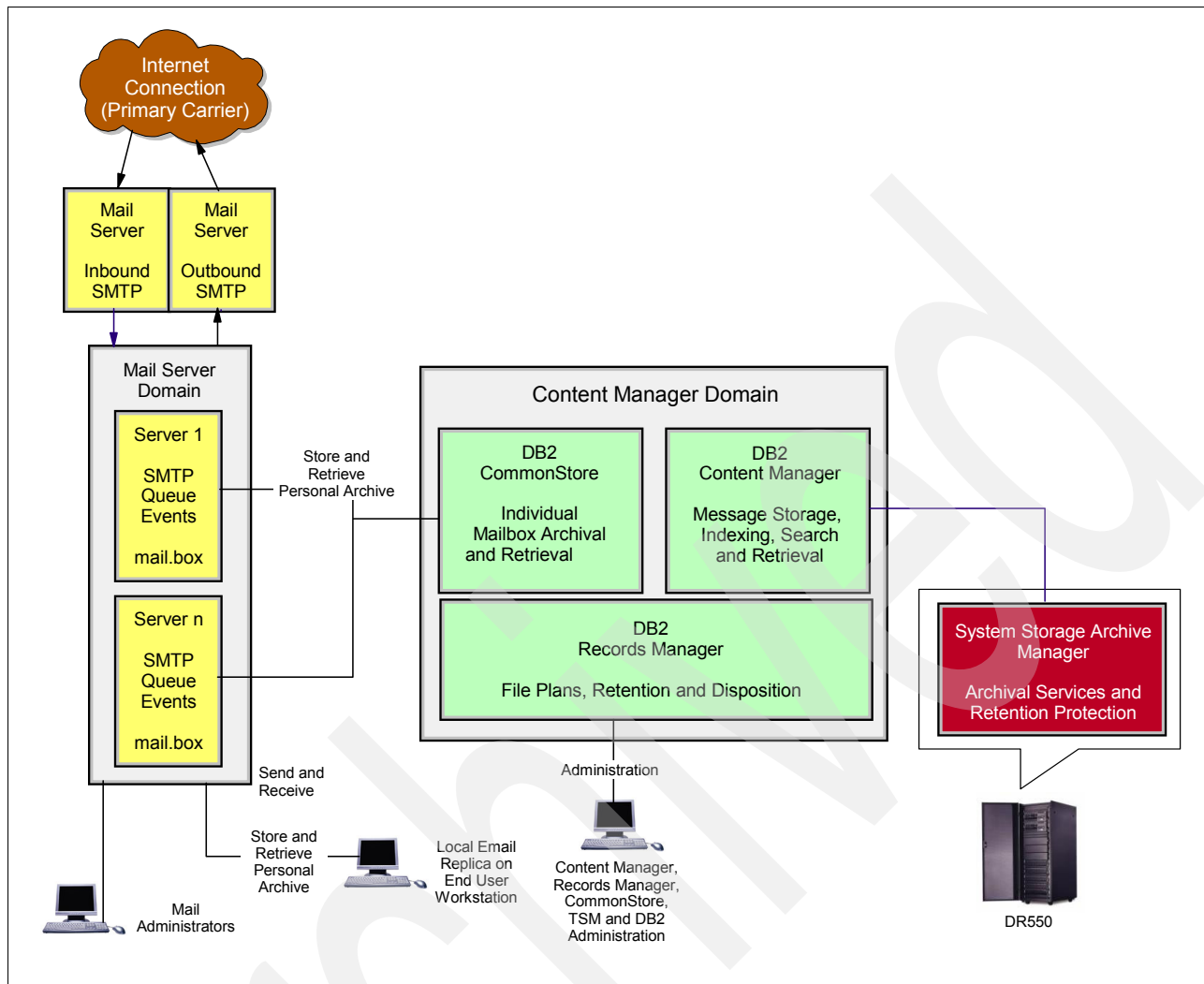
Figure 2-1   Architectural overview

Note: The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ from the current versions.

The e-mail archiving solution described here has a three-tier architecture, where an e-mail server such as Lotus Domino or Microsoft Exchange is a client application, considered as Tier 1 to this infrastructure solution. The core of the infrastructure solution consists of the middleware in Tier 2 that links the Tier 1 application to the storage in Tier 3. Major middleware used in Tier 2 consists of DB2 CommonStore, DB2 Content Manager, and DB2 Records Manager. DR550 is used as storage in Tier 3.

2.3 Component description

The following sections describe the major components used in the solution.

2.3.1 IBM DB2 Content Manager

IBM DB2 Content Manager Enterprise Edition Version 8.3 is a scalable enterprise content management (ECM) solution that enables users to leverage all of their digital information for maximum impact. From multimedia to text, DB2 Content Manager supports a range of information formats and makes content available across multiple applications and workgroups. With full-text search capabilities for both meta data and text-based documents, DB2 Content Manager allows users to easily locate pertinent information.

DB2 Content Manager, which supports a multi-tier, scalable, and distributed architecture, includes the following components that are utilized in this solution:

► Resource Managers: Specialized repositories optimized for managing the storage, retrieval, and archiving of enterprise content. The hierarchical storage management capabilities of DB2 Content Manager leverage the storage management capabilities of IBM Tivoli® Storage Manager to support a broad range of magnetic, optical, and tape devices.

► Library Server: A central source for indexing, describing, locating, organizing, and managing enterprise content.

► Client for Microsoft Windows: The desktop client that provides out-of-the-box support for high-volume, high-performance applications. Capabilities include full-text search options, full support for multi-valued attributes, and version control for annotations.

2.3.2 Content Manager Records Manager Enabler (CMRE)

In Version 8.3, IBM DB2 Content Manager provides a Content Manager Records Manager Enabler, tightly integrating DB2 Content Manager with robust, highly scalable records management tools. The integration with DB2 Records Manager ensures automatic, consistent, and accurate records declaration and classification across all forms of business information, streamlining compliance and reducing administrative costs. The Content Manager Records Manager Enabler can be used to:

► Declare e-mail messages and attachments as records with existing e-mail applications, such as Lotus Domino and Microsoft Exchange Server

► Declare corporate documents stored in Content Manager as records, and search, retrieve, view, and print record contents with existing clients, such as eClient, DB2 Content Manager, and DB2 Document Manager

► Store declared records in a centralized enterprise repository, that being Content Manager

► Use workflow user exits to customize processes that perform record declaration within DB2 Content Manager Document Routing

2.3.3 IBM DB2 CommonStore for Lotus Domino Server

DB2 CommonStore for Lotus Domino enables customers to archive e-mails, attachments, or folder content from their Lotus Domino systems to scalable and secure back-end archive management repositories, allowing an expansive range of storage options.

The features and capabilities of DB2 CommonStore are implemented through template modification so that they appear to users as Lotus Domino functions.

With DB2 CommonStore Version 8.3, IBM delivers new integration with DB2 Records Manager, enabling the capability to declare e-mail messages and attachments as records while archiving them, either with user-based selection, drag-and-drop activity, or fully automated without user involvement. With this built-in integration, users can manage, retain, and dispose of e-mails as records based on regulatory, legal, and corporate requirements, thereby improving operational efficiency while addressing compliance needs.

Some of the major capabilities delivered with DB2 CommonStore for Lotus Domino include:

► Tight integration with Lotus Domino V6 and V6.5 native archiving interfaces enables you to trigger DB2 CommonStore archiving policies through Lotus Domino V6 and V6.5 archiving requests.

► Archival can be initiated by individual users on any message or document they select or can be driven by automated pre-scheduled policies without an user involvement.

► Lotus Notes® Client provides full text indexing capability on the attributes and body of messages, with support for all types of attachments.

► More granular options for archiving support record management, including the ability to archive just the attachments, archive attachments together with the message body, archive each component as an individual item, or archive just a single item.

► Reduce storage requirements by having the same messages archived only once in the central archive repository with Single Instance Store, and compressing the content before storing it.

► DB2 CommonStore can fully exploit the security features of Lotus Domino, as well as its own features, to further refine archive security. The secure hypertext transfer protocol (HTTPS) can be used to communicate with the DB2 CommonStore server, preventing unauthorized users from accessing critical data while it is sent to or received from the server.

► Leveraging Tivoli Storage Manager, DB2 CommonStore can be integrated with enhanced storage support, including the IBM System Storage Data Retention 550 system, providing flexibility to deploy the most appropriate device based on the life cycle value of the content.

2.3.4 IBM DB2 CommonStore for Microsoft Exchange Server

DB2 CommonStore for Microsoft Exchange Server seamlessly integrates with Microsoft Exchange messaging systems to offload messages, attachments, and public folder content from the server or local PST files to a highly scalable IBM content management repository. It also provides easier migration from Microsoft Exchange 5.5 to Microsoft Exchange 2000 or 2003 Servers. Once messages are archived, users can continue to easily and securely retrieve their e-mails and attachments directly from the Microsoft Outlook client or Microsoft Outlook Web Access client – even years later.

DB2 CommonStore for Microsoft Exchange Server allows convenient, automated, rule-based archiving – for example, to store useful collections of shared information from Microsoft Exchange public folders. It also lets users archive their e-mails interactively, which provides flexibility.

Some of the major capabilities delivered with DB2 CommonStore for Microsoft Exchange Server include:

► Microsoft Exchange 5.5, 2000, and 2003 Server support can be leveraged for easier migration from Microsoft Exchange 5.5 to Microsoft Exchange 2000 or 2003 Server.

- More flexible archiving and retrieval functions, including automatic archiving of messages by age (delivery date as well as time of last modification), by mailbox size, and by user groups, and the ability to view or retrieve messages from an archive using a standard Web browser.

- Full text indexing capability from the Microsoft Outlook client on attributes and the body of messages, with support for all types of attachments.

- Space-saving new deletion options, including moving the entire message from the mailbox to archival storage without leaving a retrieval stub behind, or deleting only a space-wasting attachment and replacing it with a retrieval URL.

- Leveraging Tivoli Storage Manager, DB2 CommonStore can be integrated with enhanced storage support, including IBM System Storage Data Retention 550 systems, providing the flexibility to deploy the most appropriate device based on the life cycle value of the content.

- Enhanced security by using HTTPS protocol for secure transmission, preventing browser caching that would otherwise retain and display proprietary company information to subsequent, unauthorized users.

- Reduced storage requirements by having the same messages archived once in the central archive repository with Single Instance Store and compressing the e-mails before storing them.

2.3.5 IBM DB2 Records Manager

DB2 Records Manager is a tool for applying formal records management policies and practices to electronic and non-electronic information. When used within the context of a clear and consistent corporate policy, DB2 Records Manager can offer a number of benefits, including:

- Reduced litigation risk through structured document destruction

- Reduced operational and litigation discovery costs

- Demonstrated compliance with regulations that affect your business

- Enhanced decision-making through access to timely and relevant business information

DB2 Records Manager provides the underlying record-keeping infrastructure on a remote dedicated server. The record-keeping processes are embedded into the business application interface so that the host application can do the following:

- Declare

  Documents can be managed as records (declared) through process-driven automatic methods, or manual, user-driven methods. Once a document has been declared a record, full lifecycle management will be applied to it.

- Classify

  DB2 Records Manager assigns the appropriate retention and disposition rules to the declared record.

- Lifecycle management:

  DB2 Records Manager destroys or transfers declared records within a business application, based on the assigned rules, through consistent auditable records scheduling and disposition actions.

DB2 Records Manager applies control to declared records within the applications and does not duplicate any application functionality. Declared records are maintained within the application's repository, not removed from it. DB2 Records Manager acts as a decision

engine that lets the host application do its job, without interfering with daily business processes. DB2 Records Manager can help companies meet compliance requirements with easy integration and high performance.

The file plan, along with a retention schedule for electronic records, is the critical foundation of any structured record-keeping environment. With the sophisticated Recordkeeping Methods Modeling (RMM) of DB2 Records Manager, the file plan can be easily implemented and modified to suit business needs.

2.3.6 IBM System Storage DR550

The IBM System Storage DR550 is designed to help businesses meet the growing challenge of managing and protecting retention-managed data and other critical information assets with operational efficiency. The IBM System Storage DR550 Machine Type 2233 is available in two models: A smaller offering targeting small and medium businesses (Model DR1), and a larger offering targeting medium and large businesses (Model DR2). Information about both IBM System Storage DR550 models can be found at:

http://ibm.com/systems/storage/disk/dr/index.html

The IBM DR550 is a comprehensive, pre-configured offering designed to help organizations store, retrieve, manage, share, and secure regulated and non-regulated data.

It is designed to provide advanced storage management technology to enable the management and enforcement of data retention policies. The DR550's policy-based archive data retention capabilities are designed to support non-erasable, non-rewritable data storage, and help address the needs of regulated industries and other businesses with long-term data retention and protection requirements.

It offers archival and retention capabilities using the latest in Serial Advanced Technology Attachment (SATA) disk drives to help organizations meet new government and industry regulatory requirements. Ordered, delivered, and installed as an integrated unit, this IBM DR550 storage offering provides upgrade options for connectivity and storage capacity, helping to manage up to 224 TB of physical disk storage capacity per system. It also supports external storage devices, such as optical or tape systems, for expansion to hundreds of TB in storage capacity. Because tape has a much lower cost than disk storage, it is well suited for long-term retention of large quantities of archive information. The DR550 uses policies to manage information across both disk and tape systems, which enables it to deliver exceptionally low overall cost of ownership while also delivering outstanding performance.

Enabling companies to protect their data when it is transmitted over the network or saved to disk, encryption can provide enhanced security for businesses through 128-bit AES or 56-bit DES encryption technology. Encryption options allow the DR550 to manage encryption keys transparent to the application or allow an application to manage encryption keys externally to the DR550.

The DR550 brings a number of industry-leading capabilities to the data retention marketplace and is designed to:

► Enable management of data that has no explicit retention period, such as employee (as long as they are employed) and customer (as long as the account is open) data, through an event-based records management feature. It can protect these records from deletion until a specific event occurs.

► Enable data management on multiple tiers of storage, including disk, tape, and optical storage to provide a more cost-effective solution.

- Allow a designated object or group of objects to be protected against the normal end of life (policy expiration) process by using a deletion hold management feature. This can be very useful in the event that a record or set of records needs to be retained for legal, audit, or other reasons.
- Enforce data retention polices that maintain data as non-erasable and non-rewritable until deletion is permitted by retention policy.
- Enable users to archive and retrieve directly from or to their workstations or file servers to protected storage through the System Storage Archive Manager Client.

DR550 can also help protect customer data during disasters. IBM System Storage DR550 provides support for Metro or Global Mirroring. These features allow two real-time synchronous or asynchronous copies of the data to be maintained on DR550s in separate locations. It also provides the capability to interface with advanced WORM based tape drives and libraries to provide efficient and cost-effective copies of the data objects and the DR550 database to support off-site data storage and recovery in the event of a failure or the need to relocate to an alternate facility.

For additional information about IBM System Storage DR550, refer to *IBM System Storage DR550 V4.5 Setup and Implementation*, SG24-7091.

2.4 Coordination of retention

Now we discuss how to coordinate retention among Content Manager, Records Manager, and System Storage Archive Manager.

Content Manager (CM), Records Manager (RM), or System Storage Archive Manager (SSAM) can control the retention of an object, depending on how you configure the system. Even if you have retention policies set in RM or CM, SSAM can override them if the configuration is not done correctly. It is critical to make sure that your RM, CM, and SSAM settings are in sync, to ensure that objects are retained as you expect.

Whether SSAM or RM/CM is ultimately in control of retention depends on whether you are using chronological or event-based retention:

- With chronological retention, objects expire a defined number of days after they are archived. The retention period is fixed, and is managed by SSAM. In this mode, only the retention policies defined in SSAM are enforced and there is no coordination of retention policies between CM/RM and SSAM.
- With event-based retention, SSAM will take no action until an event is triggered by an application (such as RM/CM). What is done when the event is triggered (that is, SSAM deleting the object immediately or waiting for a defined number of days) depends on the setting of other retention-related parameters, as described in the following sections.

Table 2-1 on page 13 summarizes the SSAM archive copy group parameters that are related to retention management.

Table 2-1   SSAM archive copy group parameters

Parameter and purpose	Value for chronological retention	Value for event-based retention
RETINIT Determines when to initiate the retention period defined in RETVER attribute	CREATION (start counting when object is archived)	EVENT (start counting when event is triggered)
RETVER Number of days to retain the archived object after retention is initiated	0 to 30,000 days, or NOLIMIT	0 to 30,000 days
RETMIN Minimum number of days to retain archived object	Not applicable	n, where n represents a number of days

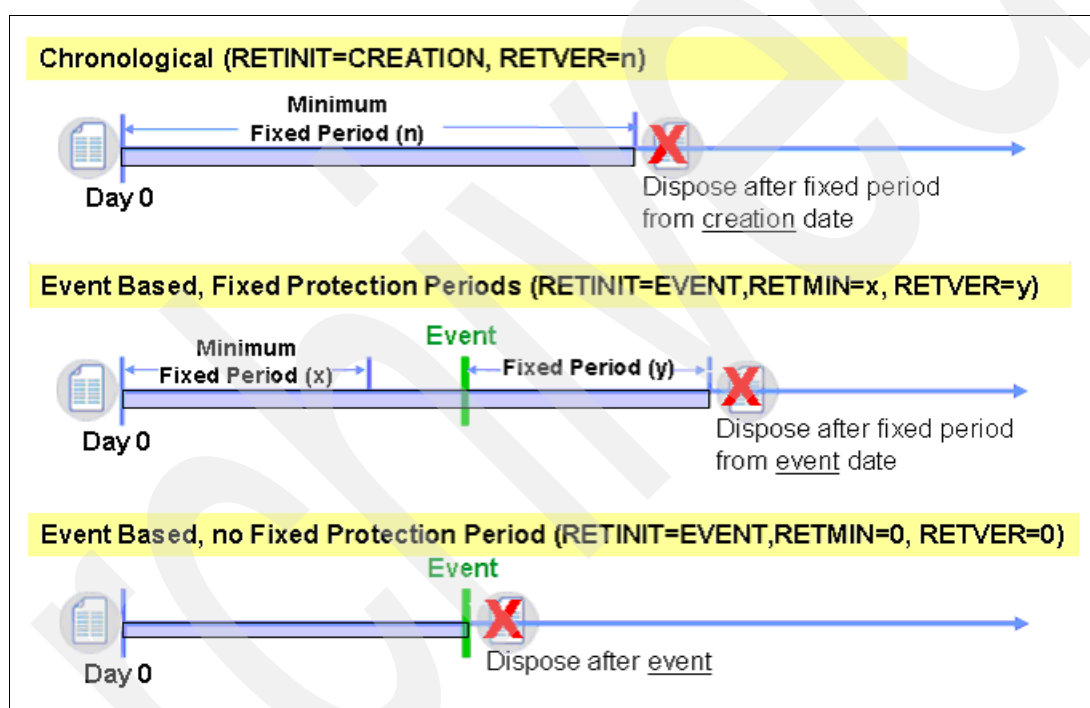Figure 2-2 illustrates these concepts.

Figure 2-2   Chronological versus event-based retention

2.4.1 Recommendation

For correct management of records, you should set RETINIT=EVENT, RETMIN=0, and RETVER=0. With this configuration, RM will be responsible for retention of objects that are declared as records, and will send the appropriate event notification to IBM Tivoli Storage Manager/SSAM (note that strictly speaking, RM notifies CM, and it is CM that actually triggers the event). For copy groups that use these settings, IBM Tivoli Storage Manager/SSAM will simply delete objects when it receives the event notification. This is what you want because RM contains the logic to ensure that retention policies are managed correctly, including management of legal holds and releases (legal holds and releases are difficult, if not impossible, to handle correctly using chronological retention).

2.4.2 Implementation

This section describes the implementation of our retention setup.

Step 1: Create the required objects in SSAM, Content Manager and Records Manager

To coordinate retention policies among CM, RM, and SSAM, different objects must be created.

In SSAM, define a policy domain, then a policy set, a management class, and finally, a copy group.

In CM Resource Manager, define the following objects in this sequence: a server definition of the SSAM, its device manager properties, and a storage class. Then define a IBM Tivoli Storage Manager volume, which will establish a real-time connection with the target management class in SSAM, followed by the creation of storage group, migration policy, and workstation collections.

In CM Library Server, define an item type and its attributes. Then define the document management relationships of the item type, which will point to the desired workstation collections in the CM Resource Manager.

Finally, in the RM, define a life cycle code that contains a retention policy. The company's file plan can be created next, and can refer to the desired lifecycle code as well as the target CM Library Server database, which contains the target item type's information.

For additional information about item types, file plans, lifecycle codes, and so forth, refer to *E-mail Archiving and Records Management Integrated Solution Guide Using IBM DB2 CommonStore and DB2 Records Manager*, SG24-6795.

Figure 2-3 on page 15, which is from *IBM System Storage DR550 V4.5 Setup and Implementation*, SG24-7091, shows how RM, CM, and IBM Tivoli Storage Manager, with an event-based retention policy, allow RM to control retention and disposition. RM determines that it is appropriate to delete data, RM orders CM to delete data, and CM issues a IBM Tivoli Storage Manager API *event*. Finally, the IBM Tivoli Storage Manager automated expiration process deletes the data in Storage Manager.
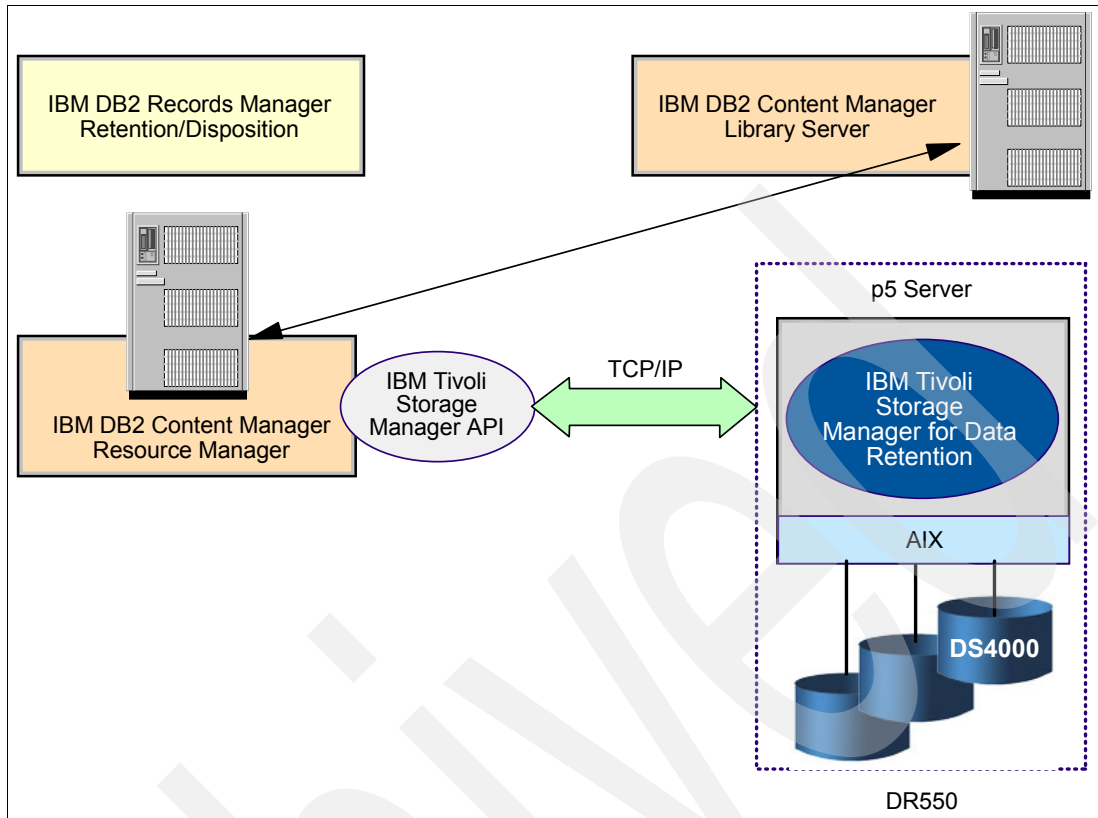
Figure 2-3   How RM controls retention and disposition

In CM, we set the migration policy in the Resource Manager so that the archived documents will be moved to the next defined tier of storage. The migration policy set in the CM is time-based. Each archived document is tied to a defined item type, which is associated with a workstation collection that has a defined migration policy. We set the migration policy in such a way that it points to a IBM Tivoli Storage Manager storage class to be the *forever* destination, which means that the defined IBM Tivoli Storage Manager storage class is the final destination of the archived (migrated) documents.

Meanwhile, in SSAM, a retention policy is defined as event-based with no fixed protection period. The retention policy is set through the creation of a copy group that points to a target destination, which is the DS4000 disk system in the DR550. The parameters set for the retention policy are given in 2.4.1, "Recommendation" on page 13.

Since the retention policy is set in SSAM to be event-based, the higher level application is responsible for notifying SSAM of the event and telling SSAM what to do with the information when the event occurs:

1. For records, the higher level application is RM. RM internally handles time, event, and event+time based retention periods, and issues an event notification (to CM, which then notifies SSAM) when all conditions of the internal retention rules are satisfied, and any holds have been released.

2. For non-records, the higher level application is CM (RM is not managing this content as records). CM should define a base retention period (simple and time-based), and notify SSAM when that period expires. Note that for e-mail archived into CM through CommonStore, e-mail Search can apply a CM hold. CM will notify SSAM of a disposition event only when its time-based retention period has expired, as well as any CM holds.

2.5 Reference implementation

To illustrate the proposed e-mail archiving and retention solution, the following infrastructure is built:

Application Layer   Consists of e-mail servers, e-mail client tool, and administration tools

Middleware Layer   Consists of DB2 CommonStore, Content Manager, and Records Manager

Storage Layer   Consists of DR550 and a tape library

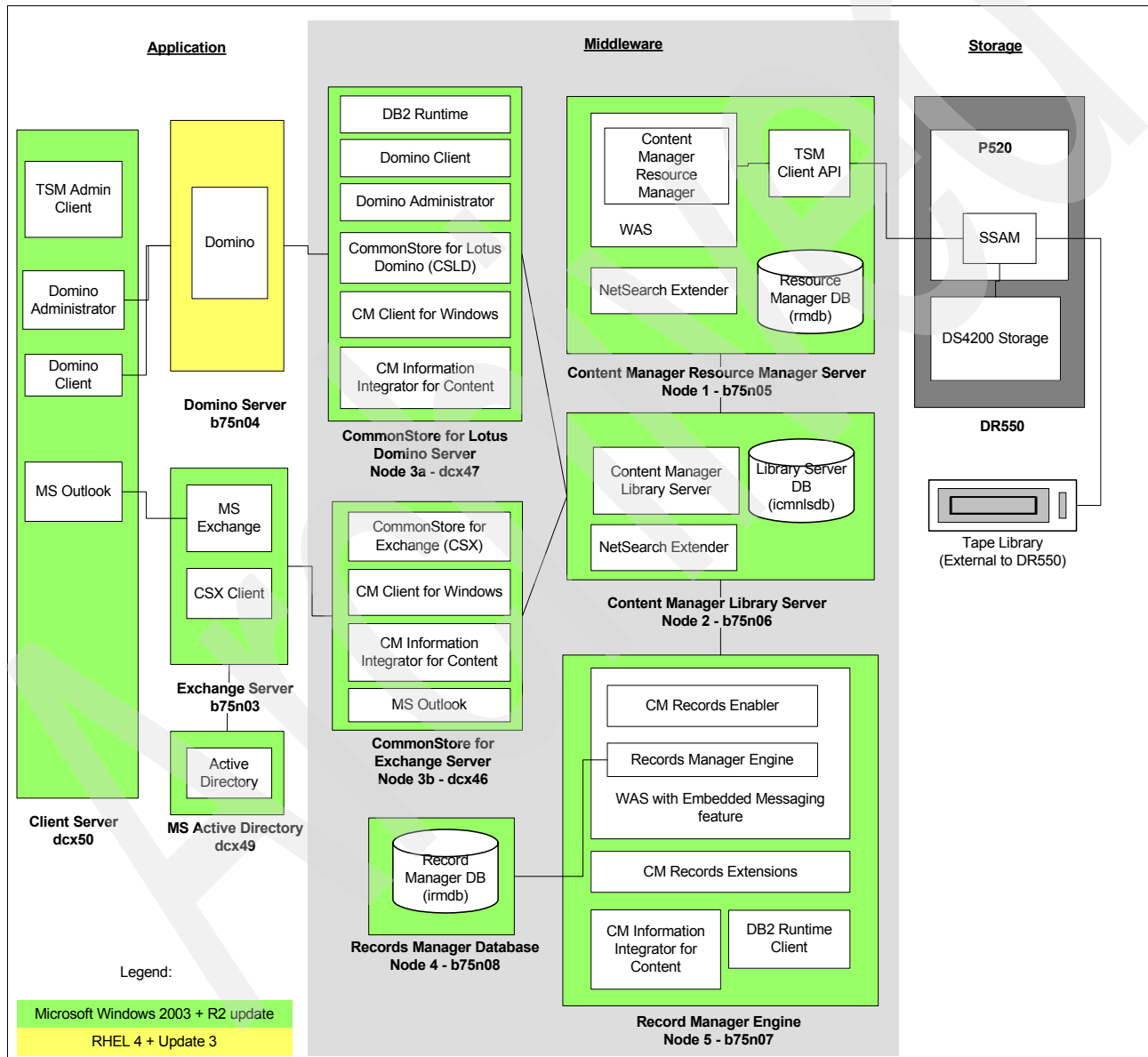Figure 2-4 shows the e-mail archiving and retention solution.

Figure 2-4   E-mail archiving and retention solution: Reference implementation

2.6 Reference implementation limitation

The infrastructure that is built for this reference implementation is mostly Microsoft Windows Server-based, except for the Lotus Domino Server running on Red Hat Enterprise Linux (RHEL) and the IBM System p server running AIX on DR550.

However, note that use of the same operating system for the e-mail server and CommonStore is recommended if there is a requirement for users to declare records from the e-mail application. Currently, there is no CommonStore and Records Manager support on Linux.

To enable users to declare records from the Lotus Domino client, there are certain shared library files that must be copied from the CommonStore machine to the Lotus Domino Server machine during the configuration. If CommonStore and Lotus Domino are running in different operating systems, as we have demonstrated in this reference implementation, the shared library files from the CommonStore machine may not be compatible with a different platform when copied over to the Lotus Domino Server machine. In this reference implementation, the shared library files from the CommonStore machine are in .dll format; these are Microsoft dynamic link library files, and are not compatible with Linux.

Therefore, in this reference implementation, users or administrators can declare records from the Content Manager Client for Windows tool and not from the Lotus Domino client, due to the compatibility issue.

Details of the integration are discussed in Chapter 4, "Implementation" on page 29.

2.7 Bill of materials

This section describes the ordering process for the individual components of the solution.

> **Note:** The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ from the current versions.

System Assurance Review

Prior to implementing this solution, a System Assurance Review (SAR) must be conducted. This will be similar in nature to the SAR that would normally be required for implementation of a DR550. Use the DR550 SAR template to conduct the SAR. A copy of the template can be found at the following URL:

http://www.ibm.com/partnerworld/pwhome.nsf/weblook/ben_trs.html

From this page, select the **Technical resources** link to sign in to PartnerWorld. To go directly to the Technical resources link, use the following URL:

http://www.ibm.com/partnerworld/pwhome.nsf/weblook/ben_detail.html?OpenDocument&type=technical&id=LLIE-6LLS4T

Hardware

The Baseline Configuration assumes a distributed multi-platform environment. With the exception of the DR550 Storage System, which is an AIX based component, the other servers can be deployed on any platform supported by the server component. This includes Windows OS and AIX.

In this publication, we describe a Microsoft Windows 2003 Standard Server platform for all components except Lotus Domino Server, which runs on Red Hat Enterprise Linux 4 with Update 3, and the AIX-based DR550 Storage System.

This build guide was created using a configuration based on the IBM BladeCenter®, IBM System x®, and DR550 technologies.

E-mail application layer

Table 2-2 through Table 2-5 on page 19 describes the e-mail application layer settings for the various servers.

Table 2-2   E-mail application layer specifications for Lotus Domino server

Lotus Domino server	
Machine type	Blade server
Model	HS40
Processor	4 x 3.0 GHz
Memory	8 GB
Drive	2 x 40 GB IDE
OS	Red Hat Enterprise Linux 4 with Update 3

Table 2-3   E-mail application layer specifications for Microsoft Exchange server

Microsoft Exchange server	
Machine type	Blade server
Model	HS40
Processor	2 x 3.0 GHz
Memory	8 GB
Drive	2 x 40 GB IDE
OS	Microsoft Windows 2003 Standard Server with Service Pack 1 or later

Table 2-4   E-mail application layer specifications for Microsoft Active Directory server

Microsoft Active Directory® server	
Machine type	IBM System x
Model	x336
Processor	2 x 3.2 GHz
Memory	2 GB
Drive	1 x 72 GB SCSI
OS	Microsoft Windows 2003 Standard Server with Service Pack 1 or later

Table 2-5   E-mail application layer specifications for the client server

Client server - e-mail client, application administration tools, and so on	
Machine type	System x
Model	x336
Processor	2 x 3.2 GHz
Memory	2 GB
Drive	1 x 72 GB SCSI
OS	Microsoft Windows 2003 Standard Server with Service Pack 1 or later

Middleware layer

Table 2-6 through Table 2-13 on page 21 describes the middleware layer specifications for the various servers.

Table 2-6   Middleware layer specifications for Content Manager Resource Manager server

Node 1: Content Manager Resource Manager server	
Machine type	Blade server
Model	HS20
Processor	2 x 3.6 GHz x 64
Memory	4 GB
Drive	2 x 36 GB SCSI
OS	Microsoft Windows 2003 Standard Server with Service Pack 1 or later

Table 2-7   Middleware layer specifications for Content Manager Library server

Node 2: Content Manager Library server	
Machine type	Blade server
Model	HS20
Processor	2 x 3.6 GHz x 64
Memory	4 GB
Drive	2 x 36 GB SCSI
OS	Microsoft Windows 2003 Standard Server with Service Pack 1 or later

Table 2-8   Middleware layer specifications for CommonStore for Lotus Domino server

Node 3a: CommonStore for Lotus Domino server	
Machine type	System x
Model	x336
Processor	2 x 3.2 GHz
Memory	2 GB
Drive	1 x 72 GB SCSI
OS	Microsoft Windows 2003 Standard Server with Service Pack 1 or later

Table 2-9   Middleware layer specifications for CommonStore for Exchange server

Node 3b: CommonStore for Exchange server	
Machine type	System x
Model	x336
Processor	2 x 3.2 GHz
Memory	2 GB
Drive	1 x 72 GB SCSI
OS	Microsoft Windows 2003 Standard Server with Service Pack 1 or later

Table 2-10   Middleware layer specifications for Records Manager Database

Node 4: Records Manager Database	
Machine type	Blade server
Model	HS20
Processor	2 x 3.6 GHz x 64
Memory	4 GB
Drive	2 x 36 GB SCSI
OS	Microsoft Windows 2003 Standard Server with Service Pack 1 or later

Table 2-11   Middleware layer specifications for Records Manager Engine

Node 5: Records Manager Engine	
Machine type	Blade server
Model	HS20
Processor	2 x 3.6 GHz x 64
Memory	4 GB
Drive	2 x 36 GB SCSI
OS	Microsoft Windows 2003 Standard Server with Service Pack 1 or later

Table 2-12   Middleware layer specifications for storage layer

Storage layer	
Machine type	IBM System Storage DR550 Machine Type 2233
Model	DR1 or DR2
Storage pools	DS4200 (8 TB of disk space) - primary storage pool, bundled and pre-configured with DR550
Server	IBM System p520
OS	AIX 5L™ V5.3

Table 2-13   Middleware layer specifications for secondary storage layer

Secondary storage layer	
Machine type	IBM System Storage Tape Library
Model	3582
Storage pools	Secondary storage pool added to the DR550

This configuration was deemed sufficient for testing and documentation purposes. However, each customer environment is different. We recommend that you work with a Product Specialist or with TechLine to properly size the server components for your customer's environment. Of course, the recommended servers must meet each solution component's minimum system requirements as published in the component's documentation. Also, keep in mind some typical customer constraints, such as standard purchasing contracts, internal IT standards for hardware, the customer's current inventory of equipment, and so on. Each of these constraints must factor into your server recommendation.

The initial storage hardware order is based on the IBM System Storage DR550, with 8 TB of disk space. In addition, we recommend adding an LTO3 or LTO4 tape autoloader. A typical configuration is shown, but you should work with a Product Specialist or TechLine when ordering, to validate that this configuration is still available and is the right configuration for your customer.

> **Note:** The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ with current versions.

Software

The following software items are described in the solution overview:

► Microsoft Windows 2003 Standard Server and Service Pack 1 or later

► Red Hat Enterprise Linux 4 and Update 3

► WebSphere® Application Server eAssembly for Windows (c53iaml.exe)

► WebSphere Application Server for Windows Fix Pack 1 (was51_fp1_win.zip)

► WebSphere Application Server for Windows Cumulative Fix Pack 10 (was511_cf10_win.zip)

► WebSphere Application Server for Windows Cumulative Fix Pack 11 (was511_cf11_win.zip)

► IBM HTTP Server 1.3.28 (part of the WebSphere Application Server eAssembly package)

► DB2 UDB ESE V8.2 for Windows (C58RSML)

► DB2 UDB ESE V8.1 for Windows Fix Pack 12 (FP12_WR21368_ESE.exe)

► DB2 Runtime Client V9.1 for Windows (C92FUML)

► DB2 Net Search Extender V8.2 for Windows (C80G3ML)

► DB2 Net Search Extender V8.2 for Windows Fix Pack 9 (db2nse822_win.exe)

► DB2 Content Manager Enterprise Edition V8.3 eAssembly for Windows (C82LFML.exe)

   – Content Manager Library Server

   – Content Manager Resource Manager

- – Content Manager Client for Windows
- ► DB2 Content Manager Enterprise Edition V8.3 Fix Pack 3 (update_win_8.3.00.300.zip)
- ► DB2 Content Manager Client for Windows V8.3 Fix Pack 1 (WinClient8.3.00.100.zip)
- ► DB2 Information Integrator for Content V8.3 for Windows (C82LLML)
- ► DB2 Information Integrator for Content V8.3 Fix Pack 3 for Windows (update_win_8.3.00.300.zip)
- ► DB2 Records Manager V4.1.1 for Windows (C80G9ML)
- ► DB2 Records Manager V4.1.2 Upgrade for Windows (IRM_4.1.2.100_WIN32.zip)
- ► Content Manager Records Manager Enabler V8.3 for Windows (C82WCML)
- ► Content Manager Records Manager Extensions V8.3 for Windows (C82WCML)
- ► IBM Tivoli Storage Manager (TSM) Client V5.5.1 For Windows 32-bit (C91IZML.exe)
- ► IBM Tivoli Storage Manager (TSM) Administration Center V5.5.1 (C86PTML)
- ► IBM Tivoli Storage Manager (TSM) Integrated Solution Console V6.0.1 (C86PQML)

If Lotus Domino is used as the mail server, the following software is needed:

- ► Lotus Domino Server V7 (c86wsna)
- ► Lotus Domino Client V7 for Windows (c88rdna)
- ► DB2 CommonStore for Lotus Domino V8.3 for Windows (C81ZMML.EXE)
- ► DB2 CommonStore for Lotus Domino V8.3 Fix Pack 1 for Windows (8.3.1-DB2-CSLD.zip)
- ► Lotus Domino Administrator V7 (c88rdna)
- ► Lotus Domino Designer (if the capability of letting e-mail users declare record is needed) (c88rdna)

If Microsoft Exchange is used as the mail server, the following software is needed:

- ► Microsoft Exchange Server
- ► DB2 CommonStore for Microsoft Exchange Server V8.3 for Windows (C81ZPML.EXE)
- ► Microsoft Outlook 2003
- ► Microsoft Active Directory

2.8 Skills

The following skills will be required in order to execute the delivery of this solution:

- ► Data classification expertise
- ► Experience with IBM storage products, including:
  - – IBM System Storage DR550
  - – IBM Tivoli Storage Manager
  - – LTO Tape
- ► Experience with enterprise content management or information lifecycle management products, including:
  - – DB2 Content Manager
  - – DB2 CommonStore
  - – DB2 Records Manager
- ► Experience with applications such as Microsoft Exchange or Lotus Domino

- ► Experience leading engagements and interacting with customers
- ► Implementation of the IBM System Storage DR550
- ► Implementation of DB2 Content Manager
- ► Implementation of DB2 CommonStore for Microsoft Exchange or Lotus Domino
- ► Implementation of DB2 Records Manager
- ► Implementation of IBM Tivoli Storage Manager
- ► Implementation of IBM LTO Tape technologies

The customer should first be lead through a data classification exercise. The output of this exercise will establish the policies that govern data retention, migration, and expiration. The next task is to complete any high-level customization and documentation of the architecture and the solution to fit the customer requirements.

2.9 Work plan

In general, the project will consist of the following high-level tasks:

1. Preparation phase
    a. Identify the data elements whose storage will be migrated to the new consolidated storage device (DR550) and managed by Content Manager.
    b. Make an inventory of the data on the existing servers and classify it using a standard taxonomy.

2. Deployment phase
    a. Perform the physical installation of the solution components:
        i.   DR550
        ii. Content Manager
        iii. Records Manager
        iv. CommonStore
        v.   LTO Tape
    b. Configure the policies that govern data movement, retention, and expiration.
    c. Perform the software installation of the Content Manager client.
    d. Skill transfer to the customer if the customer is going to be responsible for administration and operation after the implementation.

3. Migration phase
    a. Perform a full backup of all data to be migrated.
    b. Perform the initial migration.
    c. Re-assign disk resources to utilize the consolidated storage devices (DR550 and LTO tape).
    d. Test.

4. Post-migration phase
   a. Troubleshooting, if required.
   b. High-level customer training or knowledge transfer.
   c. Final documentation.
   d. Customer signoff.

3

Preparation

In this e-mail archiving and retention solution, setups for both Lotus Domino and Microsoft Exchange are shown to illustrate how the integration is done. However, in a production environment, it is likely that a customer will choose one e-mail server over another.

Depending which e-mail server the customer chooses, the following steps should be taken before performing the integration of the solution:

- ► If Lotus Domino is the target e-mail server, then:
  - Install and configure the Lotus Domino server.
  - Install the Lotus Domino Administrator client.
  - Assuming there are existing e-mail users, each user connects to the mail server through a Lotus Domino Client.
  - CommonStore for Lotus Domino should be used.
- ► If Microsoft Exchange is the target e-mail server, then:
  - Set up Microsoft Active Directory. The server must be promoted to a Domain Controller. Microsoft Windows 2003 Standard Server with Service Pack 1 is a sample platform.
  - Install and configure Microsoft Exchange server. Microsoft Exchange Server 2003 with Service Pack 2 is an example.
  - Assuming there are existing e-mail users, each user connects to the mail server through an Microsoft Outlook tool. Microsoft Outlook 2003 with Service Pack 2 is shown in this solution.
  - CommonStore for Microsoft Exchange should be used.

To leverage the use of record management, it is important that employees with expertise in the company's data classification, business policies, and government regulatory requirements should assist in determining an appropriate file plan, and data archival and retention policies.

Note: The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ with current versions.

3.1 DB2 Content Manager Enterprise Edition V8.3 prerequisites

The prerequisite products for the Resource Manager and the Library Server include the following:

► Node 1: Content Manager Resource Manager
  – DB2 UDB ESE V8.2 for Windows (C58RSML) + Fix Pack 12
  – DB2 Net Search Extender V8.2 for Windows (C80G3ML) + Fix Pack 9
  – WebSphere Application Server V5.1 + Fix Pack 1 + Fix Pack 10
► Node 2: Content Manager Library Server
  – DB2 UDB ESE V8.2 for Windows (C58RSML) + Fix Pack 12
  – DB2 Net Search Extender V8.2 for Windows (C80G3ML) + Fix Pack 9
  – Content Manager Enterprise Edition V8.3 (C82LFML) + Fix Pack 3

3.1.1 DB2 CommonStore prerequisites

Note: Choose Node 3a or Node 3b according to the type of mail server being used.

The prerequisite products for Lotus Domino and Microsoft exchange include the following:

► Node 3a: CommonStore for Lotus Domino (CSLD) Server
  – Microsoft Windows 2003 Standard Server and Service Pack 1 or later
  – DB2 Runtime Client V9.1 (C92FUML)
  – Content Manager V8.3 Client for Windows (C82LFML)
  – DB2 Content Manager Information Integrator for Content V8.3 for Windows (C82LLML) + Fix Pack 3
  – DB2 CommonStore Lotus Domino V8.3 for Windows + Fix Pack 1
  – Lotus Domino Administrator V7
► Node 3b: CommonStore for Microsoft Exchange (CSX) Server
  – Microsoft Windows 2003 Standard Server and Service Pack 1 or later
  – CommonStore for Microsoft Exchange V8.3 base + Fix Pack 1 = CSX V8.3.1
  – DB2 Content Manager Information Integrator for Content V8.3 for Windows (C82LLML) + Fix Pack 3
  – DB2 Runtime Client V9.1 (C92FUML)
  – Content Manager V8.3 Client for Windows (C82LFML)

3.1.2 Records Manager prerequisites

Note: We recommend that the Records Manager Database and Records Manager Engine be installed on separate machines.

The Record Manager prerequisite products include the following:

► Node 4: Records Manager DB
  – DB2 UDB ESE V8.2 for Windows (C58RSML) + Fix Pack 12
  – DB2 Records Manager (database only) V4.1.1 (C80G9ML) + V4.1.2 Upgrade
► Node 5: Records Manager Engine
  – WebSphere Application Server V5.1 for Windows (C53IAML) + Fix Pack 1 + Cumulative Fix Pack 11 = WebSphere Application Server V5.1.1.11
  – DB2 Runtime Client V9.1 (C92FUML)
  – DB2 Records Manager (engine only) V4.1.1 (C80G9ML) + V4.1.2 Upgrade

3.1.3 Content Manager Records Enabler prerequisites

The prerequisites for the Records Enabler include the following:

► Node 5: Records Manager Enabler
  – WebSphere Application Server (with Embedded Messaging) V5.1 for Windows (C53IAML) + Fix Pack 1 + Cumulative Fix Pack 11 = WebSphere Application Server V5.1.1.11
  – DB2 Content Manager Information Integrator for Content V8.3 for Windows (C82LLML) + Fix Pack 3 (CM Connector only)
  – DB2 Runtime Client V9.1 (C92FUML)
  – Content Manager Records Manager Enabler V8.3 for Windows (C82WCML)
► Node 5: Records Manager Extensions

> **Note:** Records Manager Extension must be deployed on the same WebSphere Application Server as the Records Manager engine.

  – WebSphere Application Server V5.1 for Windows (C53IAML) + Fix Pack 1 + Cumulative Fix Pack11 = WebSphere Application Server V5.1.1.11
  – Content Manager Records Manager Enabler V8.3 for Windows (C82WCML)

3.1.4 DR550 prerequisites

The IBM System Storage DR550 V4.5 or higher is a pre-configured Machine Type that includes:

► An IBM System p server that runs AIX.
► A DS4000 mid-range disk system.
► System Storage Archive Manager (SSAM) software installed and pre-configured to use the DS4000 as primary storage pool.

Ensure that the DR550 is accessible to the Ethernet network prior to the integration of this solution.

4

Implementation

This chapter ties all the parts together. In it, we describe our architectural decisions and the implementation steps that we follow.

29

4.1 Architectural decisions

This solution was built to model a production environment. The following architectural decisions were made.

Decision 1

All Content Manager, CommonStore, and Records Manager components could have been installed on a single server, but this is not the optimal solution. Content Manager consists of two major components: Library Server and Resource Manager. A Library Server handles all the queries to the Resource Manager, which tracks all the physical storage of the archived documents. Splitting these two components onto separate servers is recommended for larger installations; placing both on the same server would be fine for smaller installations.

Decision 2

The Records Manager Database and Records Manager Engine should be installed on separate machines according to the product installation guidelines for Records Manager.

Decision 3

Records Manager Extensions must be deployed on the same WebSphere Application Server as the Records Manager engine according to the product installation guidelines for Records Manager Extensions.

Decision 4

In this reference implementation, Records Manager and Content Manager Resource Manager are installed on different machines to show distinct functionality of these servers. It is possible to place both Records Manager and Content Manager Resource Manager on the same machine because both of them are WebSphere Application Server-based applications. However, the embedded messaging feature of WebSphere Application Server must be installed because it is a prerequisite for Records Manager.

Decision 5

Add a tape library as the second-tier storage. The tape library was added to show the end-to-end scenario of how data can be migrated from online or near-line storage to offline storage.

4.2 Component implementation sequence

The installation sequence is important in this solution because each major component has its own prerequisites and installation requirements, as described in Chapter 3, "Preparation" on page 25. This section outlines a recommended implementation sequence for deploying the reference implementation described in 2.5, "Reference implementation" on page 16.

> **Note:** The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ with current versions.

Step 1: Set up Content Manager on Node 1 and Node 2

For additional details about installing and configuring Content Manager, refer to *Content Manager Implementation and Migration Cookbook*, SG24-7051.

► On Node 1: Install the Content Manager Resource Manager Base Installation.

Assuming that Microsoft Windows 2003 Standard Server with Service Pack 1 or later is installed, then install and validate the following software in this sequence:

a. DB2 ESE V8.2.

b. DB2 ESE V8.1 Fix Pack 12.

c. DB2 Net Search Extender V8.2.

d. DB2 Net Search Extender V8.2 Fix Pack 9.

e. WebSphere Application Server V5.1.

f. WebSphere Application Server V5.1 Fix Pack 1.

g. WebSphere Application Server V5.1 Cumulative Fix Pack 10.

h. IBM Content Manager V8.3 with the **Resource Manager Database** and **Resource Manager Application** options selected.

► On Node 2: Set up Content Manager Library Server Base Installation.

Assuming that Microsoft Windows 2003 Standard Server with Service Pack 1 is installed, install and validate the following software in this sequence:

a. DB2 ESE V8.2.

b. DB2 ESE V8.1 Fix Pack 12.

c. DB2 Net Search Extender V8.2.

d. DB2 Net Search Extender V8.2 Fix Pack 9.

e. IBM Content Manager V8.3 with the **Library Server** option selected

► On Node 1 and Node 2: Set up a Database Connection between Resource Manager on Node 1 and Library Server on Node 2:

a. On Node 1 (where the Resource Manager is installed): Catalog the Library Server database, icmnlsdb.

b. On Node 2 (where the Library Server is installed): Catalog the Resource Manager database, rmdb.

► On Node 1 and Node 2: Validate Content Manager Base Installation:

a. On Node 1: Run the validation program.

b. On Node 2: Run the validation program.

► On Node 1: Complete the Installation of Content Manager Resource Manager:

a. Install IBM Content Manager V8.3 Fix Pack 3.

b. From the Windows Services, change ICM Migrator to run automatically. This is required to move e-mail from IBM Content Manager into SSAM.

c. Install IBM Tivoli Storage Manager Client V5.3.4.

► On Node 2: Complete the installation of Content Manager Library Server:

a. Install IBM Content Manager V8.3 Fix Pack 3.

► On Node 1 and Node 2: Validate Content Manager Fix Pack installation:

a. On Node 1: Re-run the validation program.

a. On Node 2: Re-run the validation program.

Choose Step 2a or 2b according to the type of mail server in your existing environment.

For additional details about installing and configuring CommonStore, refer to *IBM Content Manager CommonStore Version 7 for Domino, Exchange, and SAP*, SG24-6405.

Step 2a: Set up CommonStore for Lotus Domino (CSLD) on Node 3a if Lotus Domino is the e-mail server being used

Note: The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ from the current versions.

If users need to declare records from the e-mail application, CSLD and the Lotus Domino Server must be on the same operating system.

In this reference implementation, Microsoft Windows 2003 Standard Server with Service Pack 1 is installed. Install and validate the following software:

► IBM CommonStore Lotus Domino V8.3
► IBM CommonStore Lotus Domino V8.3 Fix Pack 1
► IBM Content Manager Information Integrator for Content V8.3
► IBM Content Manager Information Integrator for Content V8.3 Fix Pack 3
► DB2 Runtime Client V9.1
► Content Manager Client for Windows V8.3
► Lotus Domino Administrator V7

Step 2b: Set up CommonStore for Microsoft Exchange (CSX) on Node 3b if Microsoft Exchange is the e-mail server being used

In this reference implementation, Microsoft Windows 2003 Standard Server with Service Pack 1 is installed. Install and validate the following software:

► IBM Content Manager Information Integrator for Content V8.3
► IBM Content Manager Information Integrator for Content V8.3 Fix Pack 3
► IBM Content Manager Client for Windows V8.3
► DB2 Runtime Client V9.1
► IBM CommonStore for Microsoft Exchange V8.3
► IBM CommonStore for Microsoft Exchange V8.3 Fix Pack 1
► Microsoft Outlook 2003

Step 3: Set up IBM Records Manager and Records Manager Enabler (CMRE) on Node 4 and Node 5

Note: The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ from the current versions.

For more information about installing and configuring Records Manager and CMRE, refer to *E-mail Archiving and Records Management Integrated Solution Guide Using IBM DB2 CommonStore and DB2 Records Manager*, SG24-6795.

To view a demo of installing the Records Manager, go to:

http://www.ibm.com/developerworks/db2/library/techarticle/dm-0509kernek

Microsoft Windows 2003 Standard Server with Service Pack 1 should be installed.

► On Node 4 and Node 5: Install prerequisite components.

– On Node 4, install these components in this sequence:

i. DB2 ESE V8.2.

ii. DB2 ESE V8.1 Fix Pack 12.

– On Node 5, install these components in this sequence:

i. DB2 Runtime Client V9.1.

ii. IBM Content Manager Information Integrator for Content V8.3.

iii. IBM Content Manager Information Integrator for Content V8.3 Fix Pack 3.

iv. WebSphere Application Server V5.1.

v. WebSphere Application Server V5.1 Fix Pack 1.

vi. WebSphere Application Server V5.1 Cumulative Fix Pack 10.

► On Node 4 and Node 5: Prepare to install IBM Records Manager Database and Engine:

i. Create the Connection Factories Authentication User irmwas on Node 5.

ii. Create the Records Manager Administrator User ID irmadmin on Node 4.

iii. Set the environment variable "DB2UNIVERSAL_JDBC_DRIVER_PATH" in WebSphere Application Server on Node 5.

► On Node 4: Install IBM Records Manager database:

i. Install IBM Records Manager (database only) V4.1.1.

ii. Upgrade the database to IBM Records Manager V4.1.2.

► On Node 5: Install and configure IBM Records Manager Engine:

i. Install IBM Records Manager V4.1.1 (base).

ii. Upgrade IBM Records Manager to V4.1.2.

iii. Configure Records Manager Engine.

► On Node 5: Install IBM Content Manager Records Manager Enabler (CMRE) and IBM Content Manager Records Manager Extensions:

i. Install IBM Content Manager Records Manager Enabler V8.3.

ii. Install IBM Content Manager Records Manager Extensions V8.3.

Choose Step 4a or 4b according to the type of mail server in your existing environment.

Step 4a: Integrate CommonStore for Lotus Domino with Content Manager

The configuration steps include:

► Create an archive user for CSLD in Content Manager.
► Define the e-mail item type in Content Manager.
► Configure the ArchPro environment.

Step 4b: Integrate CommonStore for Microsoft Exchange with Content Manager

The configuration steps include:

► Create a Content Manager archive user ID for CSX.
► Create e-mail archive attributes and item type.
► Validate the connection of Content Manager and CommonStore.
► Configure the ArchPro environment.
► Set up CommonStore tasks.
► Implement Windows services.

Step 5: Integrate Content Manager with SSAM on DR550

> **Note:** The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ from the current versions.

For additional information about IBM Tivoli Storage Manager, refer to *Certification Study Guide: IBM Tivoli Storage Manager (ITSM) Version 5.2*, SG24-6693.

The configuration steps include:

► Define retention policies in SSAM.
► Register the Content Manager clients to SSAM.
► Configure Content Manager to use SSAM on DR550, including defining migration policies in Content Manager.
► Validate the integration of Content Manager and IBM Tivoli Storage Manager.

Step 6: Integrate Records Manager with Content Manager

> **Note:** The illustrations in this publication were created with prior versions of most of the software products involved. The general approach and architecture of the solution remain valid. However, some of the detailed installation and configuration steps might differ from the current versions.

The configuration steps include:

► Define a sample file plan in Records Manager.
► Create a lifecycle code in Records Manager.
► Configure Content Manager Records Manager Enabler (CMRE).

► Validate the integration of Records Manager and Content Manager.

Step 7: Set up a tape library with DR550 if a secondary storage pool is desired

The configuration steps include:

► Define the tape library as a storage pool in SSAM.

► Configure the primary storage pool to have a next tier storage pool using the tape library.

► Set up schedules for migration from the primary storage pool to the secondary storage pool.

4.3 Standard deployment

Integrate the archival and retention solution into an existing e-mail infrastructure.

Integrate with Lotus Domino

1. Create the CSLD user in Lotus Domino.

2. Copy template files and sign them.

3. Create a Configuration database and a Jobs database.

4. Create a profile for the archive task.

5. Create a profile for the retrieve task.

6. Configure the CommonStore task environment.

7. Start the CSLD tasks.

Integrate with Microsoft Exchange

1. Extend the Active Directory schema.

2. Set up an organization forms library.

3. Grant users access to the public folders.

End-to-end testing

Perform the following basic verification tests:

► Microsoft Exchange archiving and retrieving scenarios verification tests:

   a. Manual archive for e-mail from Microsoft Outlook.

   b. Manual retrieval of archived e-mail from Microsoft Outlook.

► Lotus Domino archiving and retrieving scenarios verification tests:

   a. Manual archive for e-mail from Lotus Domino Client.

   b. Manual retrieve of archived E-mail Lotus Domino Client.

► Verification tests for Content Manager:

Using the Content Manager Client for Windows, you can search Content Manager to verify that your e-mails have been successfully archived.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this Redpaper.

IBM Redbooks

For information about ordering these publications, see "How to get IBM Redbooks". Note that some of the documents referenced here may be available in softcopy only.

► *E-mail Archiving and Records Management Integrated Solution Guide Using IBM DB2 CommonStore and DB2 Records Manager*, SG24-6795

► *IBM System Storage DR550 V4.5 Setup and Implementation*, SG24-7091

Online resources

These Web sites are also relevant as further information sources:

► IBM Solutions Builder Express "*eMail Archiving and Records Compliance Foundation*", found at:

http://sbe-ic.austin.ibm.com:8888/help/index.jsp

How to get IBM Redbooks

You can search for, view, or download Redbooks, Redpapers, Hints and Tips, draft publications and Additional materials, as well as order hardcopy Redbooks or CD-ROMs, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

37

IBM Enterprise Content Management and DR550 for E-mail Archiving and Records Management Overview

Redpaper™

Enterprise Content Management solution components

Planning and implementation

E-mail server integration

This IBM Redpaper publication provides an installation and configuration guide for an e-mail archiving reference implementation. This solution leverages the use of DR550 and major IBM software components, such as Content Manager, CommonStore, and Records Manager. The objective is to create and validate the architecture that enables the use of DR550 to support an end to end Information Lifecycle Management (ILM) scenario by demonstrating, with Lotus Domino or Microsoft Exchange, e-mail archiving and retention scenarios.

This document is intended for IT professionals who are responsible for building and implementing the environment associated with ILM deployments.

REDP-4284-01