

ERA System Design Review

Day Three

May 11, 2005

ERA SDR - DAY THREE

Preservation Design

May 10, 2005

Preservation – Agenda

Description of Functionality

Key Requirements

Functional Architecture

Service Design

Design Highlights / Trades

Physical Design

RID Discussion

Preservation Description

Preservation provides

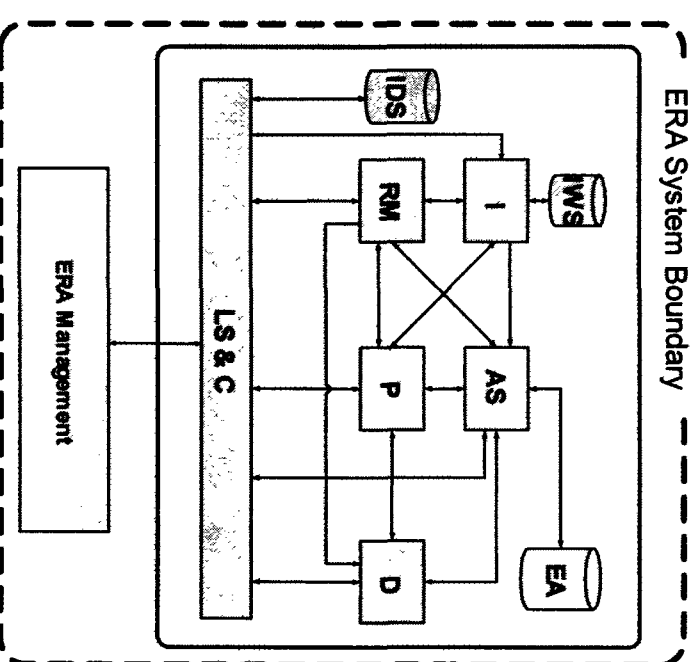
- Services necessary to manage the preservation of the electronic records to ensure their continued existence, accessibility, and authenticity over time**
- Management functionality for Preservation Assessments, Preservation and Service Level plans, Data Type Descriptors, and Digital Adaptation Descriptors**

Preservation Driving Requirements

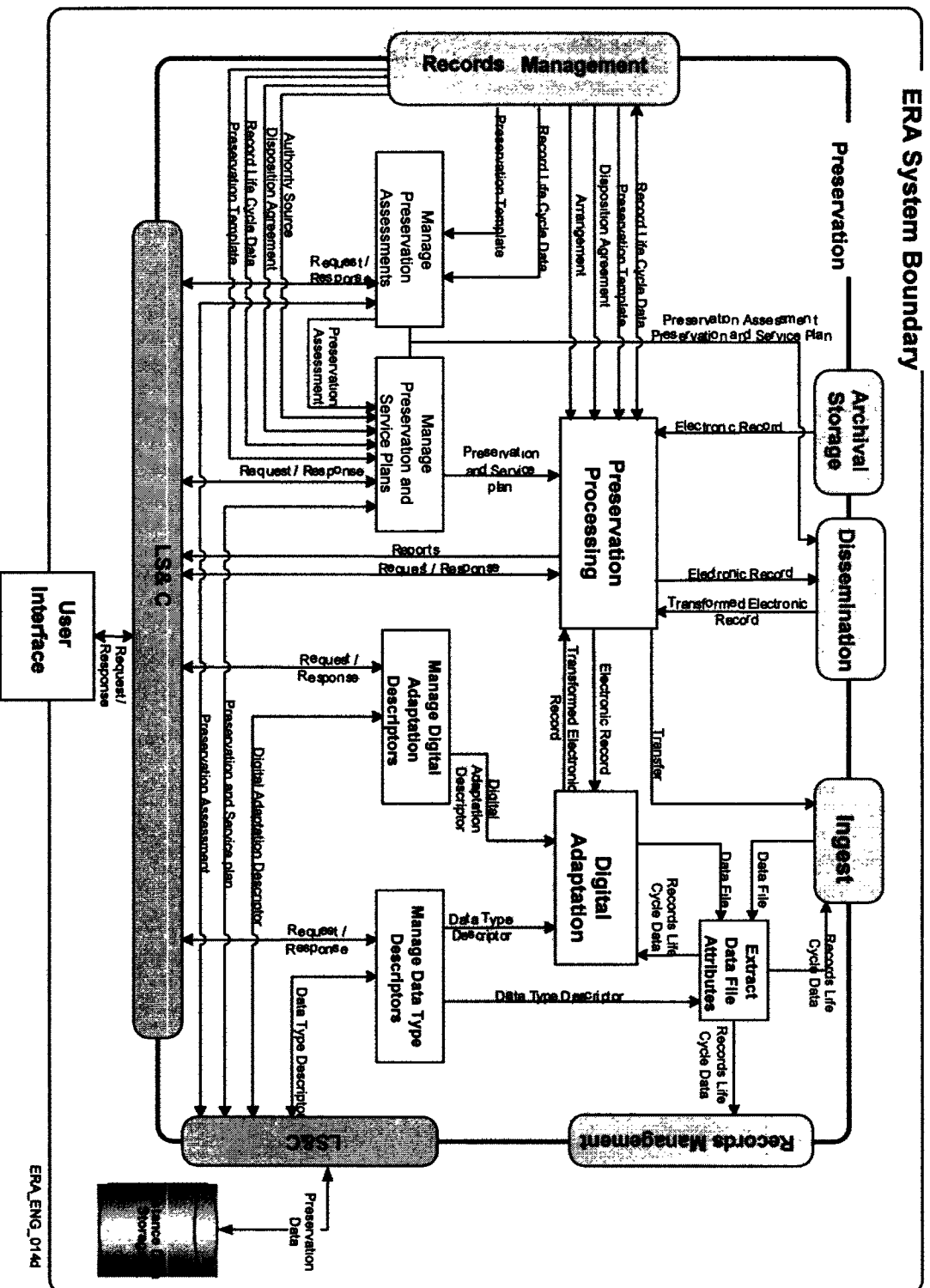
LM8.1 – Centralized facility for preservation processing services

LM8.5, LM8.6, LM8.26 – Extensible digital adaptation framework

LM8.1.1, LM8.9 – Centralized facility for preservation assessments and planning



Preservation Functional Architecture



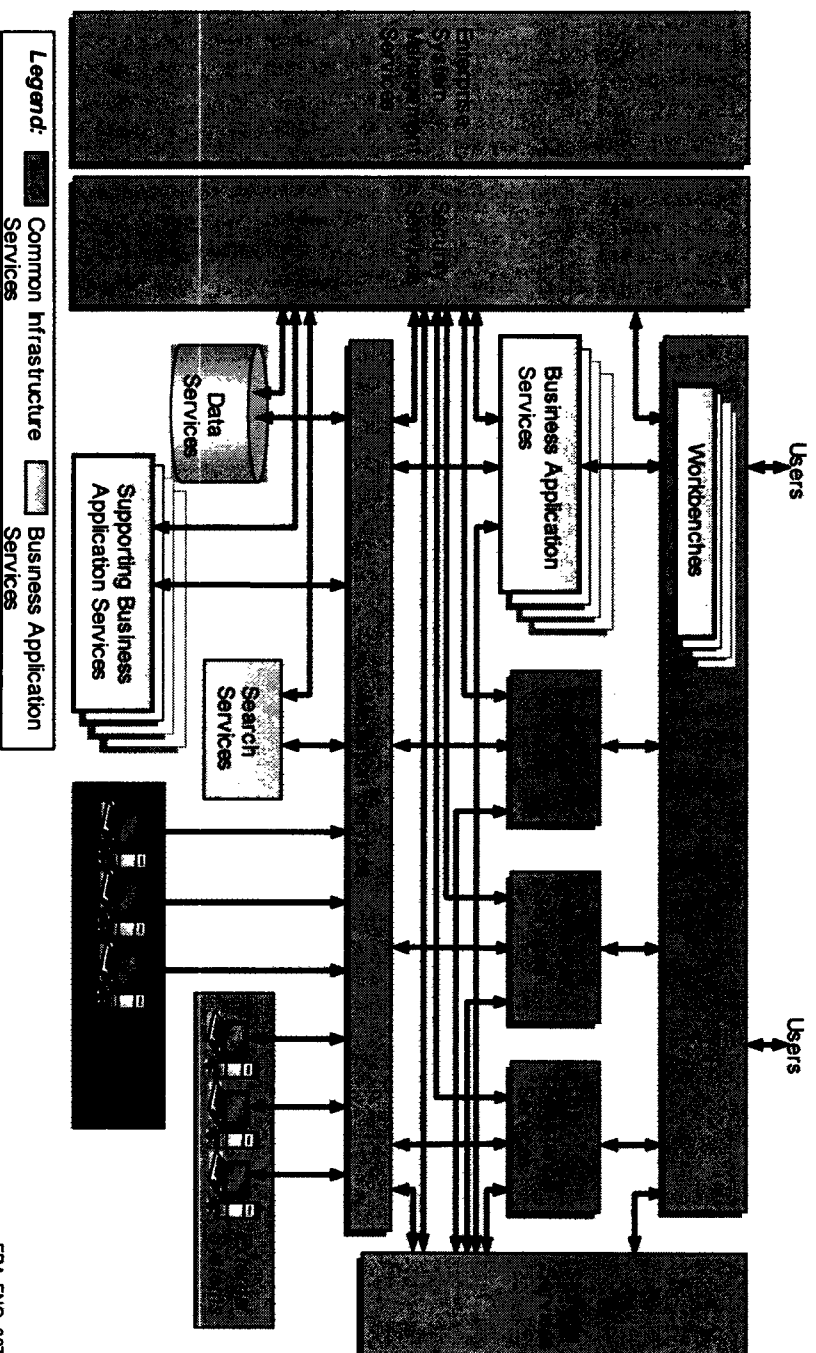
ERA_ENG_0141

- Key Features:**
- Digital Adaptation Registry
 - Data Type Registry
 - Deterministic Processing
 - Digital Adaptation Framework

Preservation within SOA

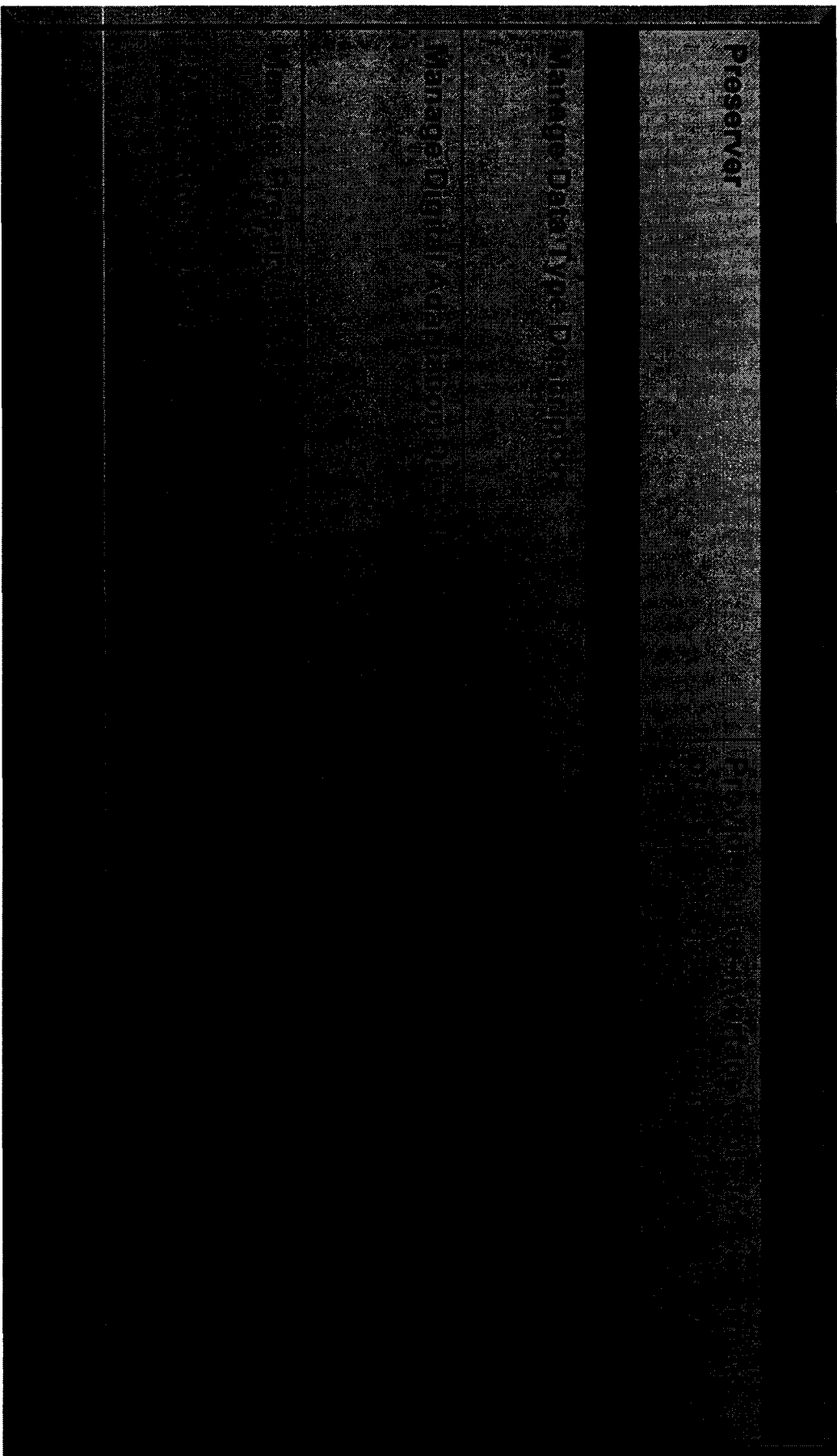
Preservation includes

- Workbenches
- Business Application Services
- Supporting Business Application Services

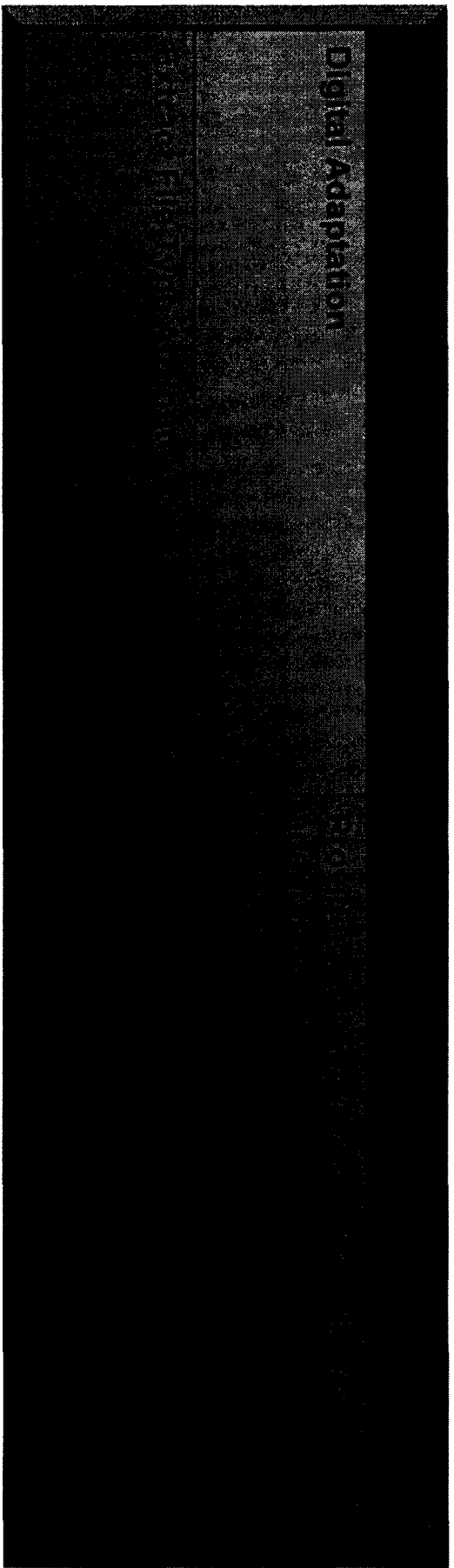


ERA_ENG_007c

Preservation Services



Preservation Services



Preservation Design Highlights

Preservation Planning

- Planning Based on Record Types and not File Types
 - Ensures consistent treatment of similar records
- Preservation Objective Model
 - Facilitates creation of deterministic preservation plans, which can be executed automatically
- Relationship to Disposition Agreement
 - Integrates preservation planning into the life cycle management of the record

Preservation Design Highlights

Preservation and Service Level Plans

- Preservation Planning**
 - **Allows preserver to decide whether records should undergo preservation processing**
- Service Level Planning**
 - **Allows preserver to define the access requirements**
- Essential Characteristics Requirements**
 - **Allows preserver to specify the essential characteristics for the records and their relative importance**
- Authenticity Requirements**
 - **Allows preserver to specify the authenticity requirements for the records**

Represented by a Preservation Objective Model

Preservation Design Highlights

Preservation Processing

–Frameworks & Registries

- Flexible and extensible framework and registries allows new data types and digital adaptation processors to be added over time

–Preservation Objective Model

- Describes the capabilities of a digital adaptation processor

–Persistence Approach

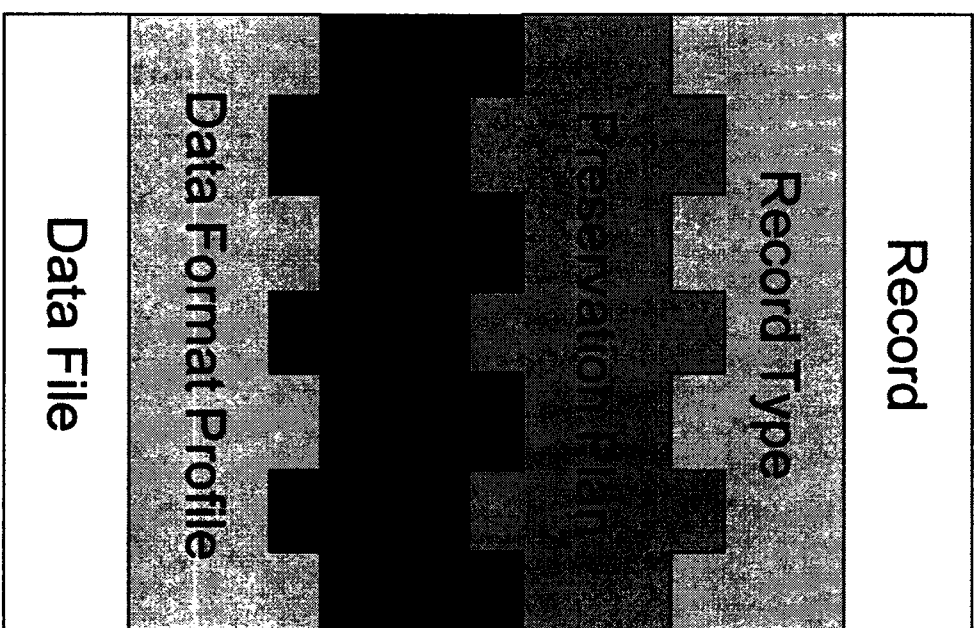
- Digital adaptation processors evaluated within context of an established intellectual framework as described in the LM Team’s “Persistent Object Format” white paper

–Authenticity Approach

- Authenticity evaluated within context of an established intellectual framework as described in the LM Team’s “Authenticity” white paper

Preservation Planning & Processing combined provide input into investment and research decisions for supporting new functionality over time

Preservation Objective Model

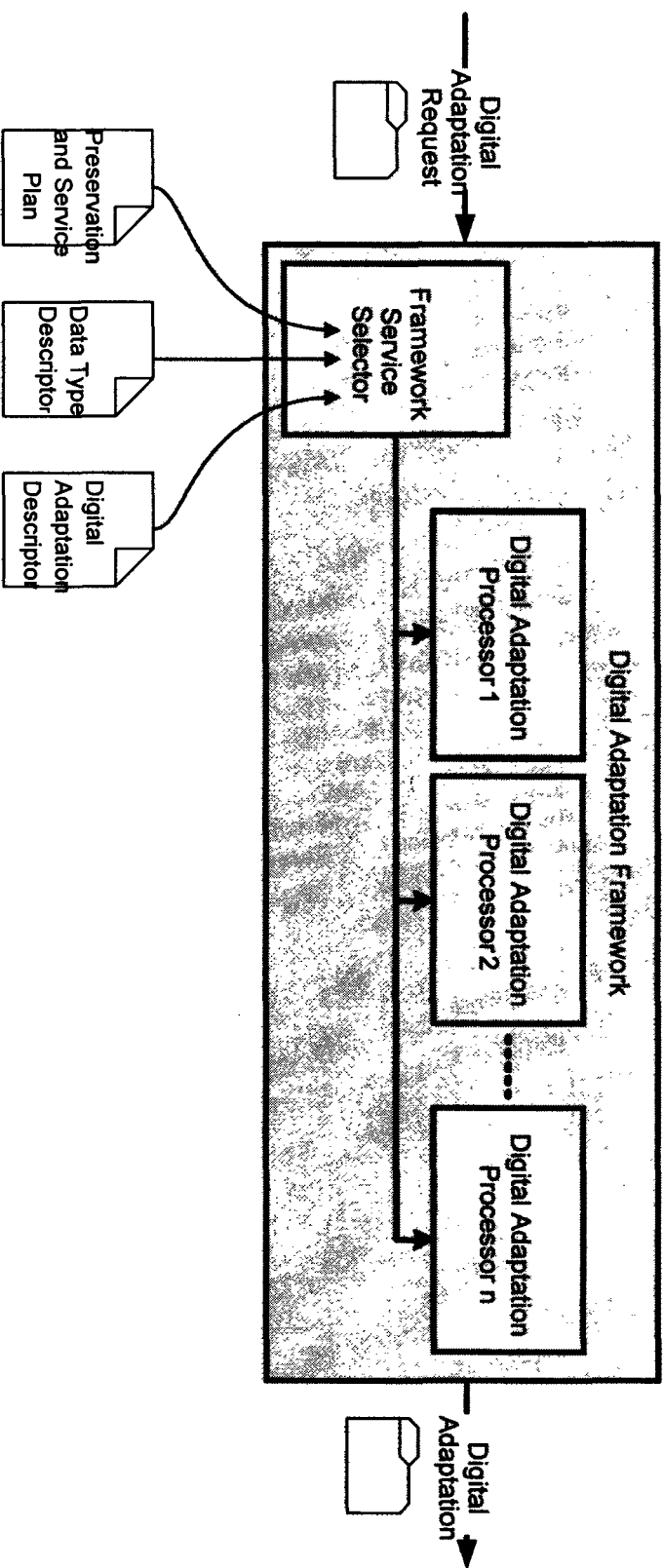


Based on requirements

Based on capability

Preservation Design Details

Digital Adaptation Framework



Key Features:

- Extensible to support new Data Types
- Extensible to support new Digital Adaptation Processors
- Supports best-fit adaptation based on preservation objectives

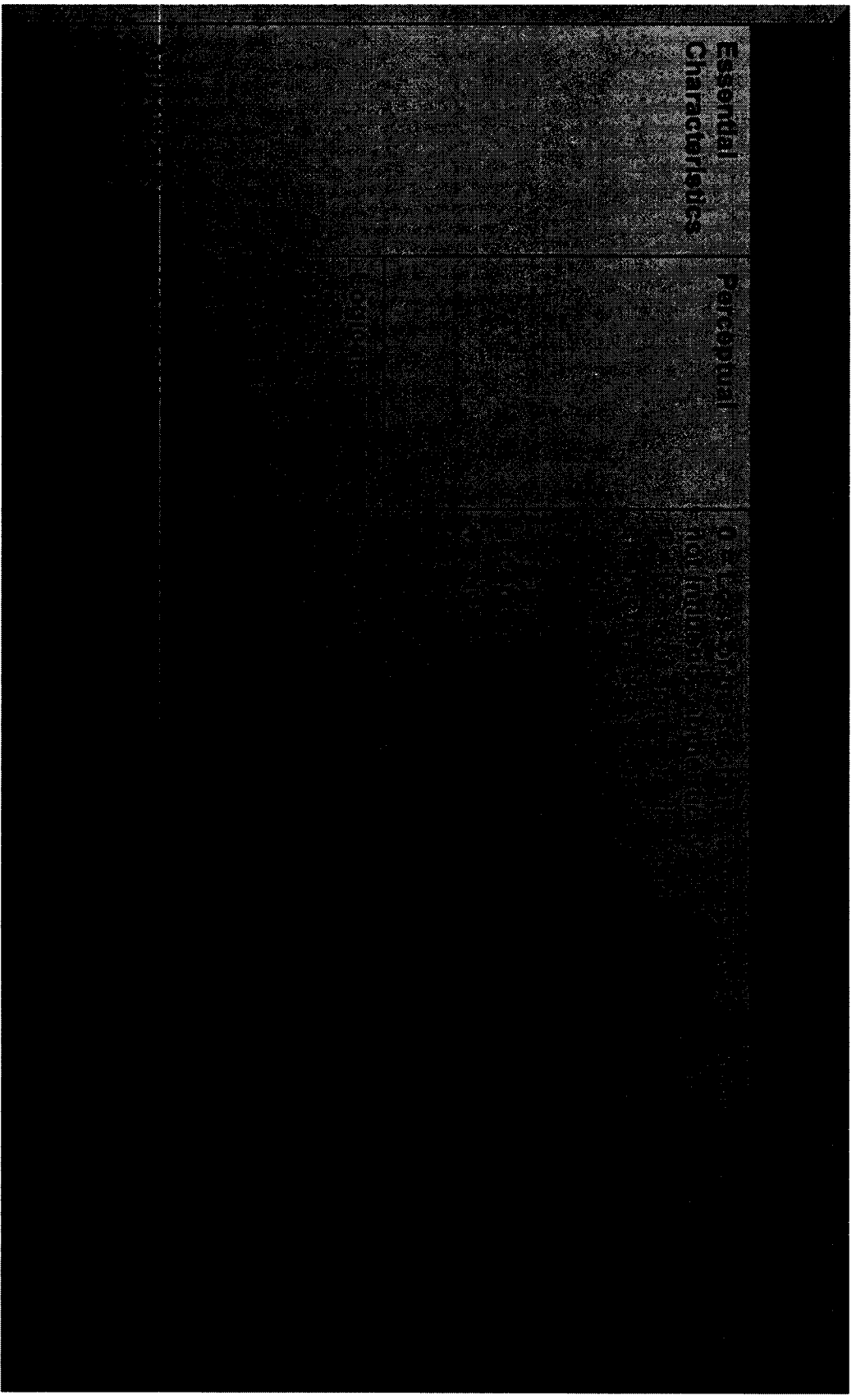
Preservation Design Details - Example

Preservation and Service Plan

Service	Location	Dependencies
[Redacted Content]		

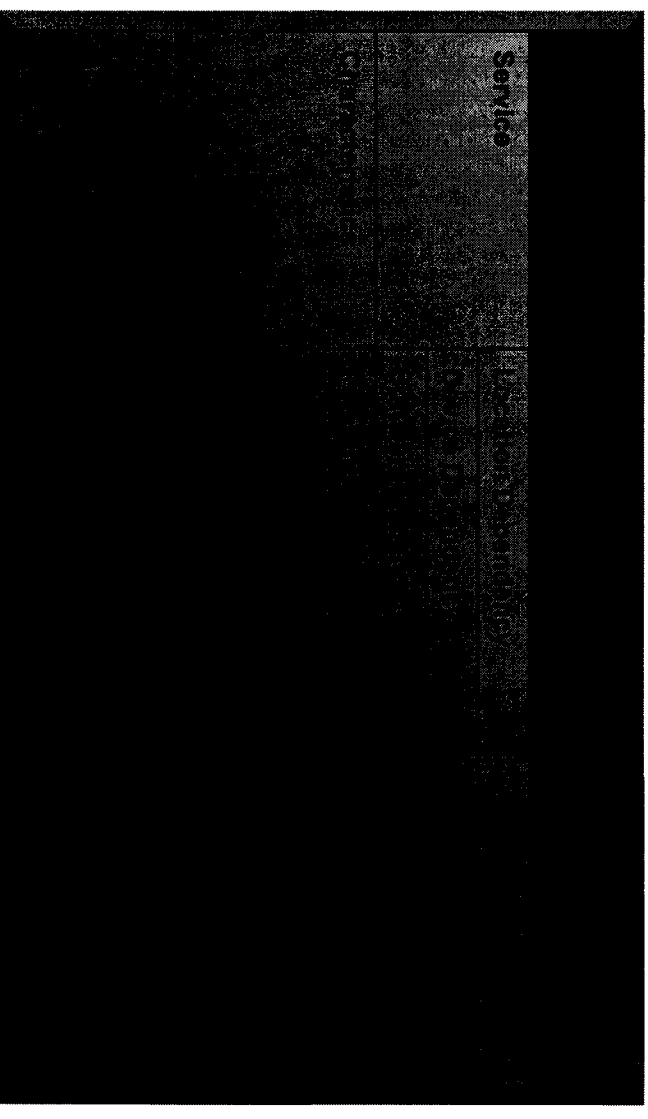
Preservation Design Details - Example

Preservation and Service Plan (continued)

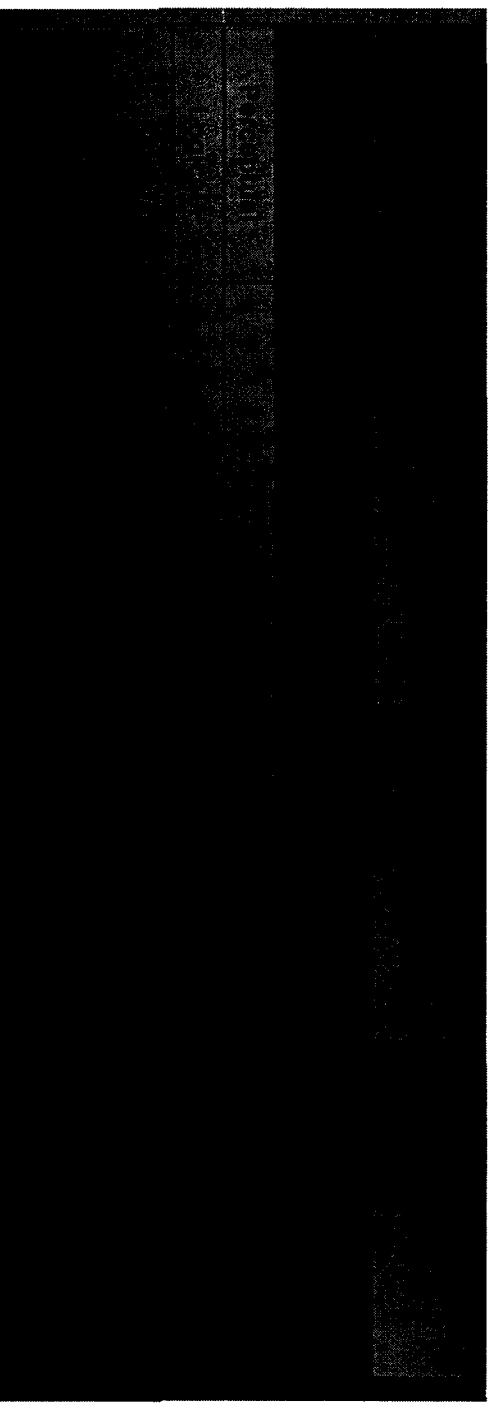
Essential Characteristics	Perceptual	Functional
		

Preservation Design Details - Example

**Example Preservation
Objective Model for a Digital
Adaptation Processor**



**Example of
Preservation
Matching for Two
Digital Adaptation
Descriptors**



Preservation Design Trades

Investigated whether existing COTS products could implement digital adaptation

- Answer really depends on a case-by-case basis, and will change over time**
- Framework allows both developed and COTS digital adaptation processors to be incorporated into ERA seamlessly**

Preservation Physical Design

Preservation services are implemented on the System/Business Applications VLAN, which is described in the Local Services & Control Design charts

Services developed as J2EE Web Services, with support from the common COTS-based infrastructure in LS&C

Preservation - Conclusions

Provide Frameworks and Registries

- Ensure durable design
- Enable collaborative development of digital adaptations

Developed Preservation Objective Model

- Provides deterministic Preservation Planning and Processing
- Captures archival judgment and policy decisions

Developed White Papers

- Provide intellectual foundation for the design

Preservation RIDs

RID-LMC00118 Preservation Functional Architecture

RID-LMC00119 Preservation

RID-LMC00120 Taxonomy of Essential Characteristics

BREAK

ERA SDR - DAY THREE

Archival Storage Design

May 11, 2005

Archival Storage Agenda

Key Requirements

Functional Architecture

Service Design

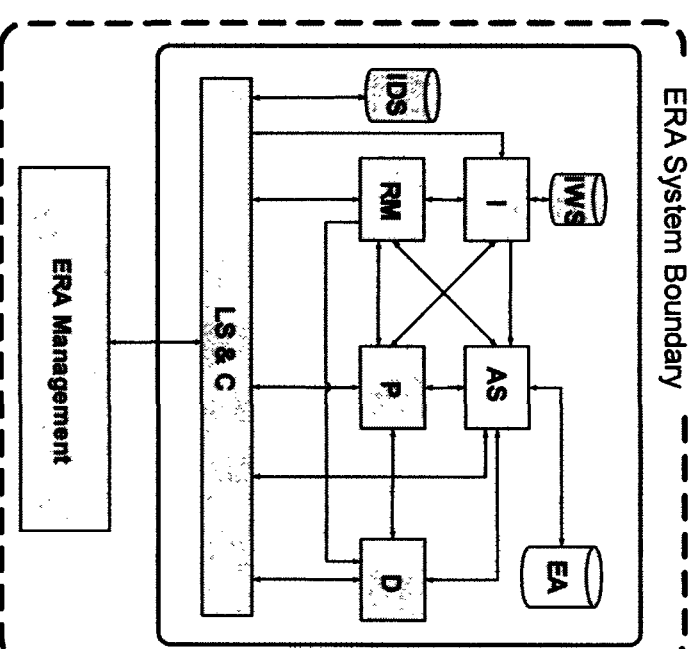
Physical Design

Initial Product Selections

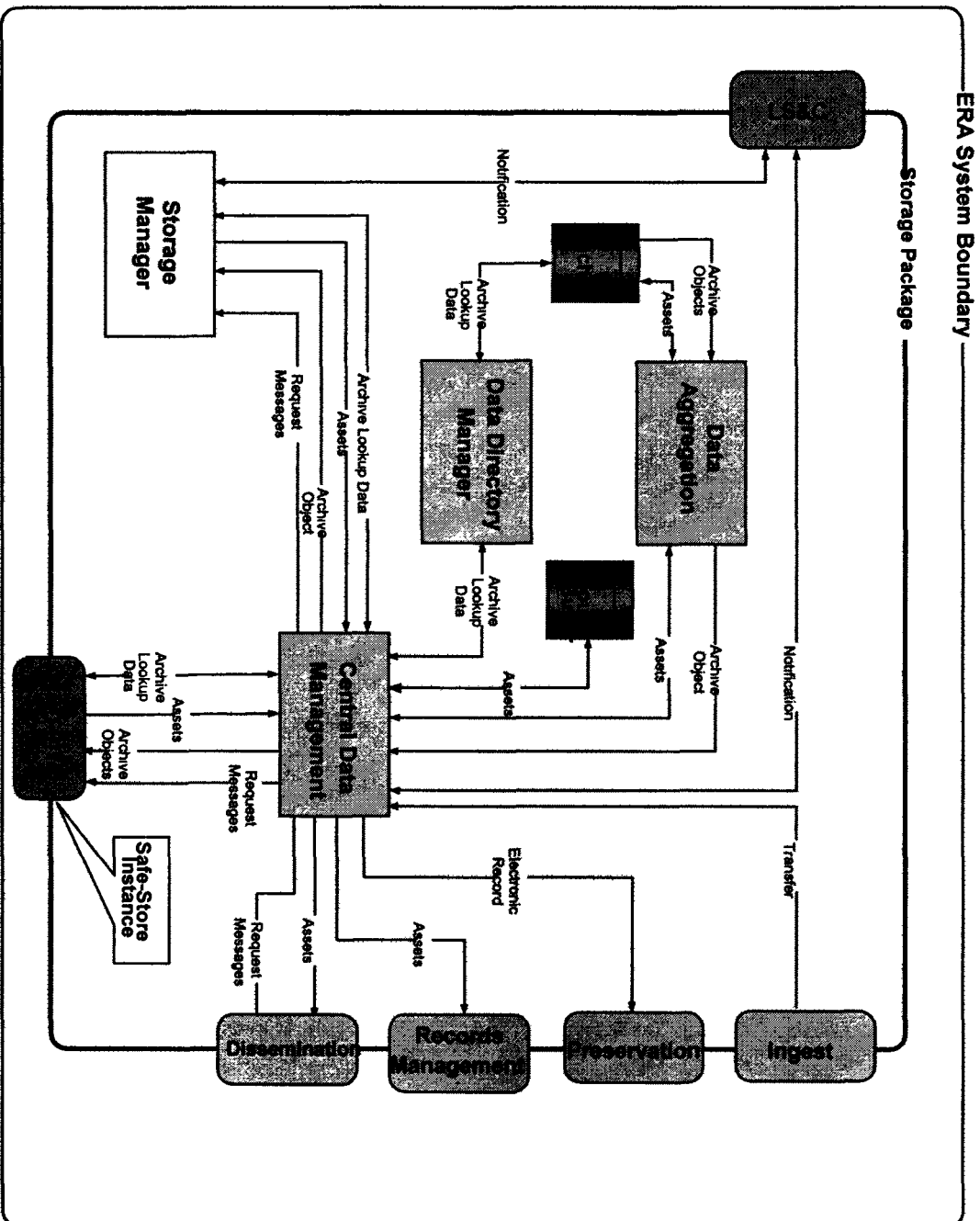
RID Discussion

Archival Storage Key Requirements

- LM10.1.4 – Active Safe-Store architecture
- LM10.2.4 – Virtual file system interface
- LM11.3 – Storage management application that abstracts removable media libraries
- LM12.1 – Solution that is not storage media or supplier dependent
- LM31.2, LM31.3 – Scalable and extensible solution



Archival Storage Functional Architecture

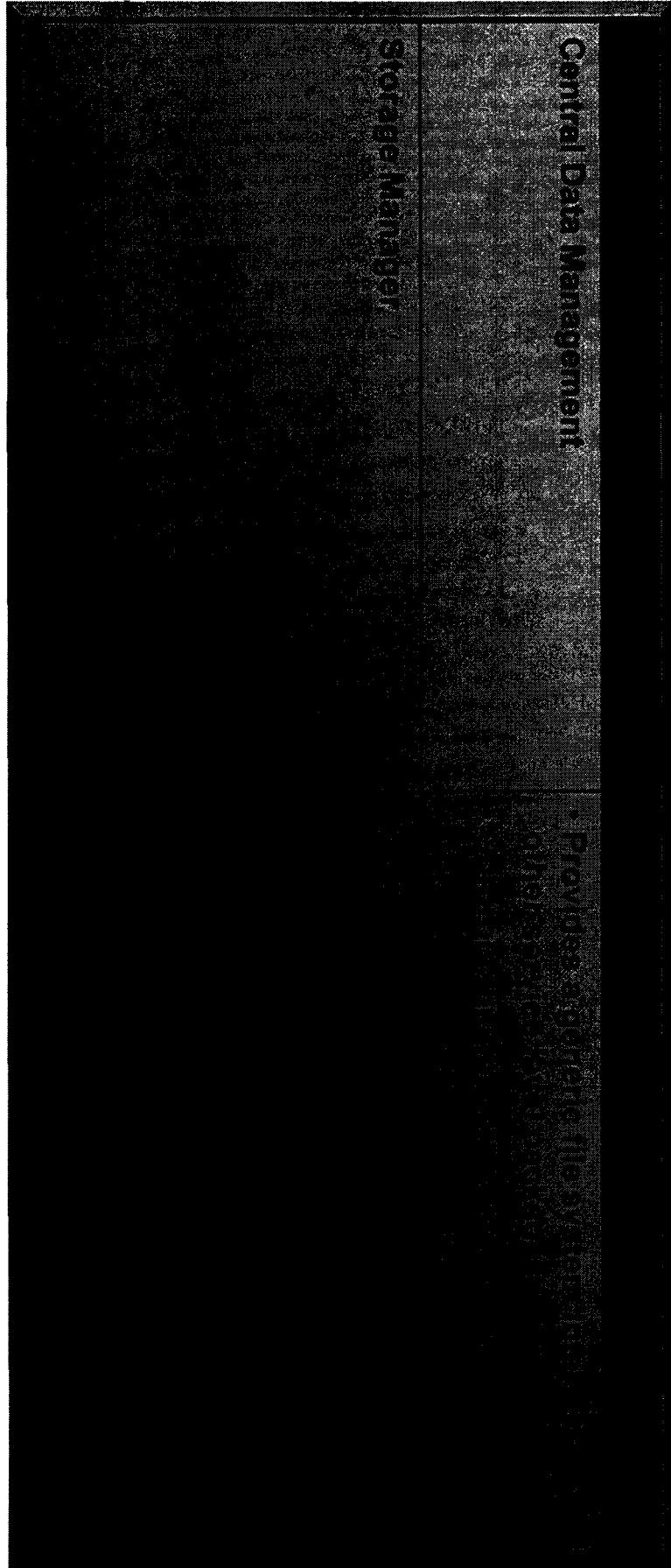


ERA_ENG_015b

Key Features:

- Active Safe-Store
- Virtual File System Interface
- Media Abstraction
- Scalable and Extensible

Archival Storage Services



Central Data Management

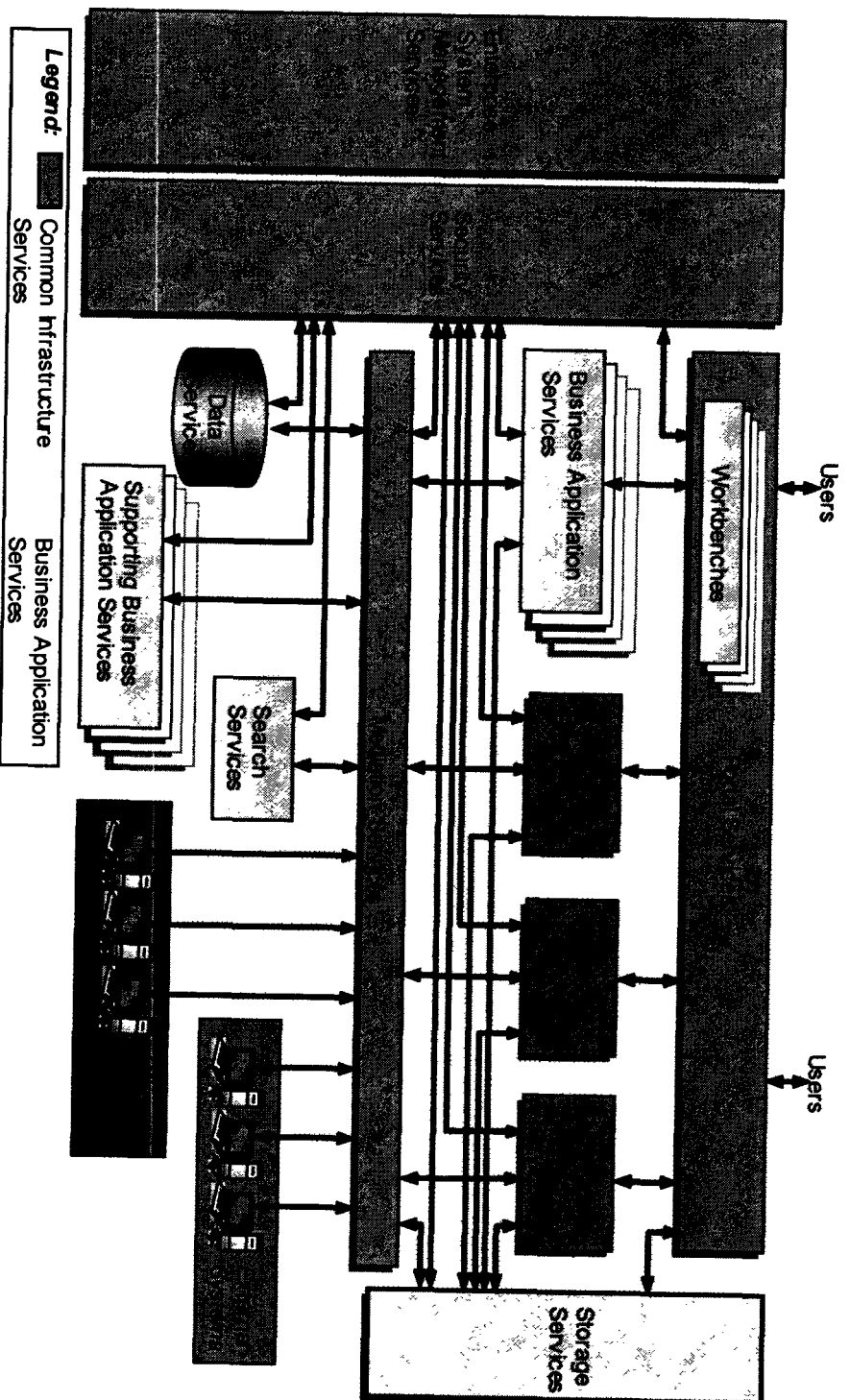
Storage Manager

Provides support for the system

Archival Storage within SOA

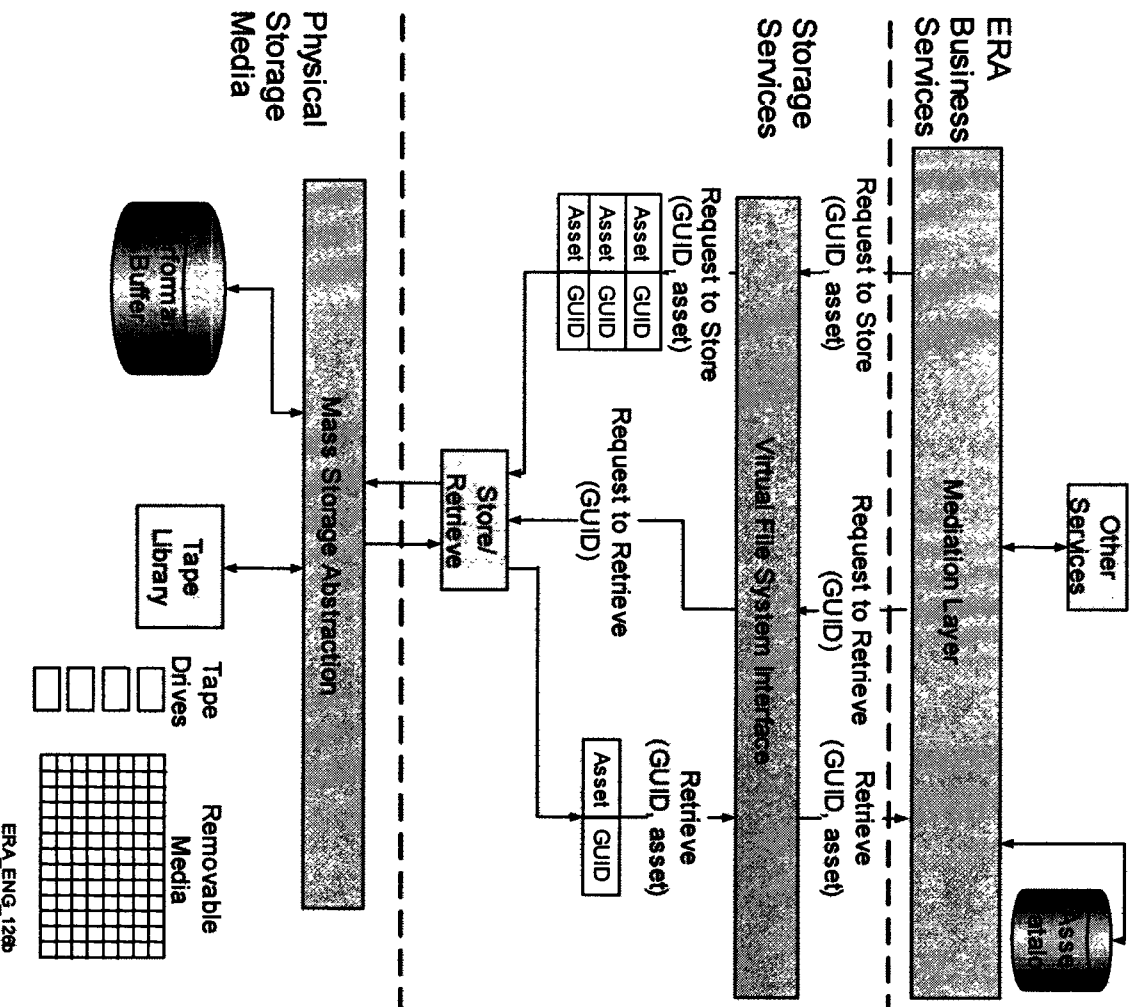
Archival Storage includes

– Storage Services (archival component)



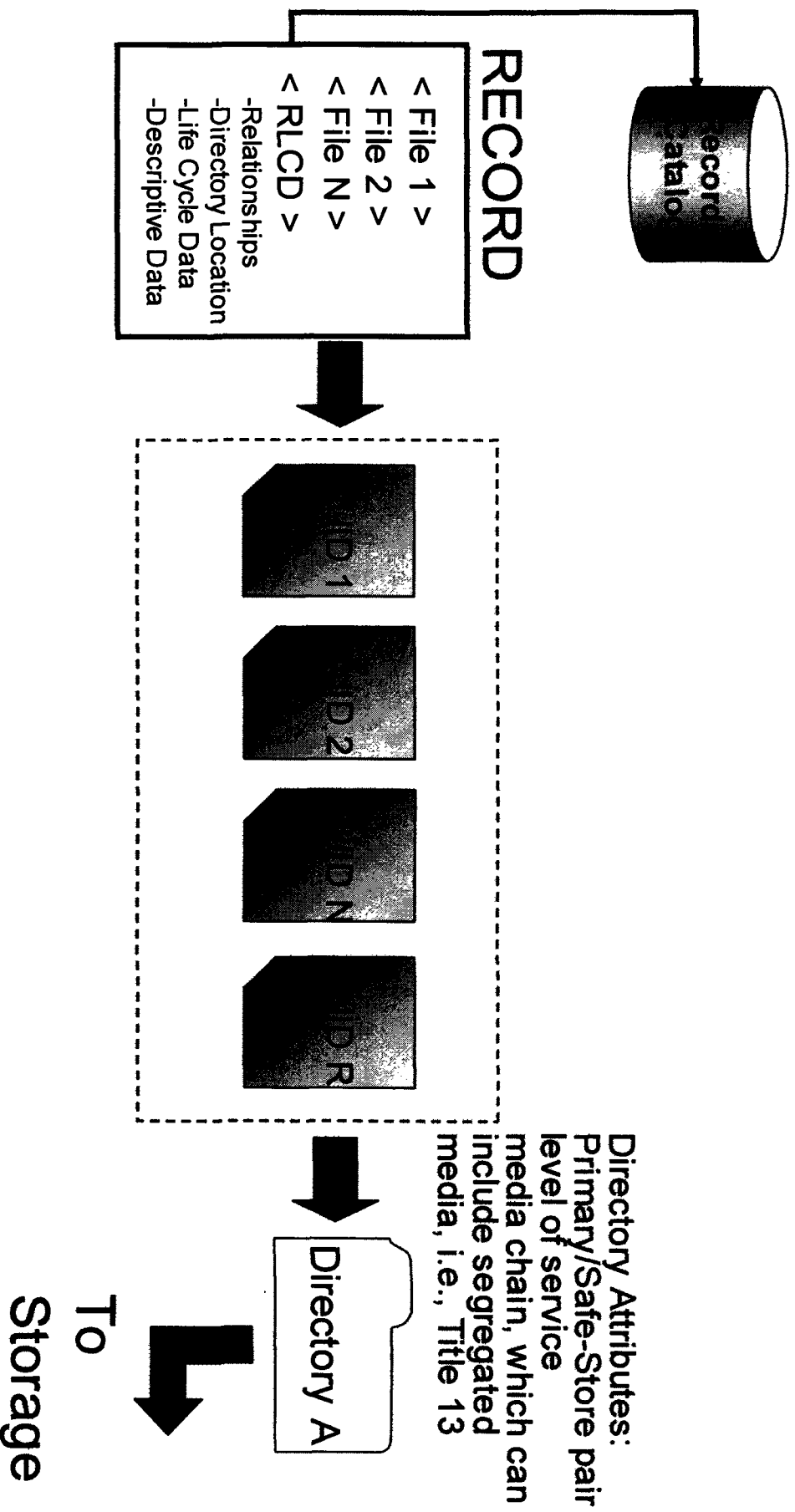
ERA_ENG_007c

Archival Storage Design Details: Functional Architecture

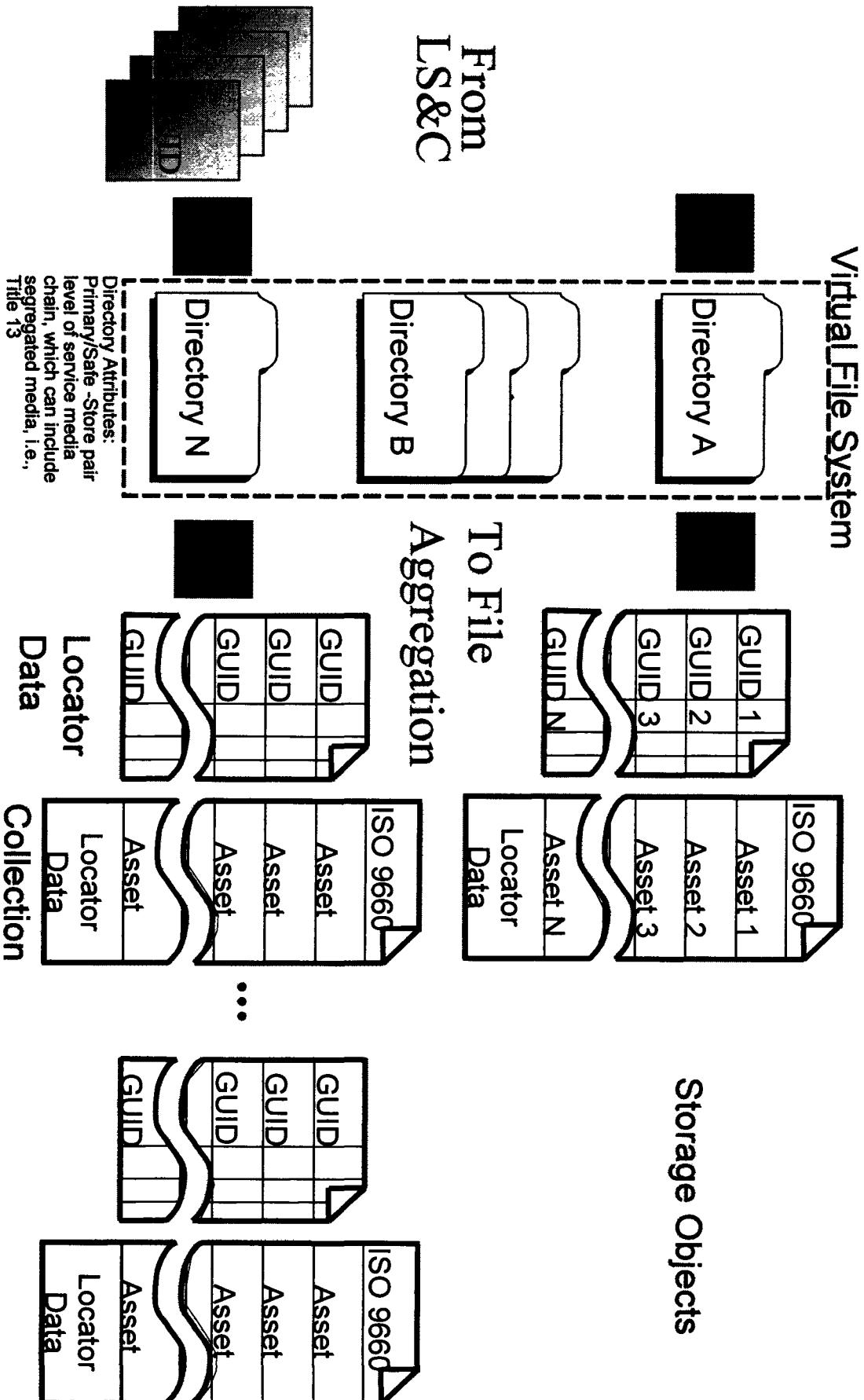


- Key Features:**
- Active Safe-Store
 - Virtual File System Interface
 - Media Abstraction
 - Scalable and Extensible

Archival Storage Design Details: Record Creation

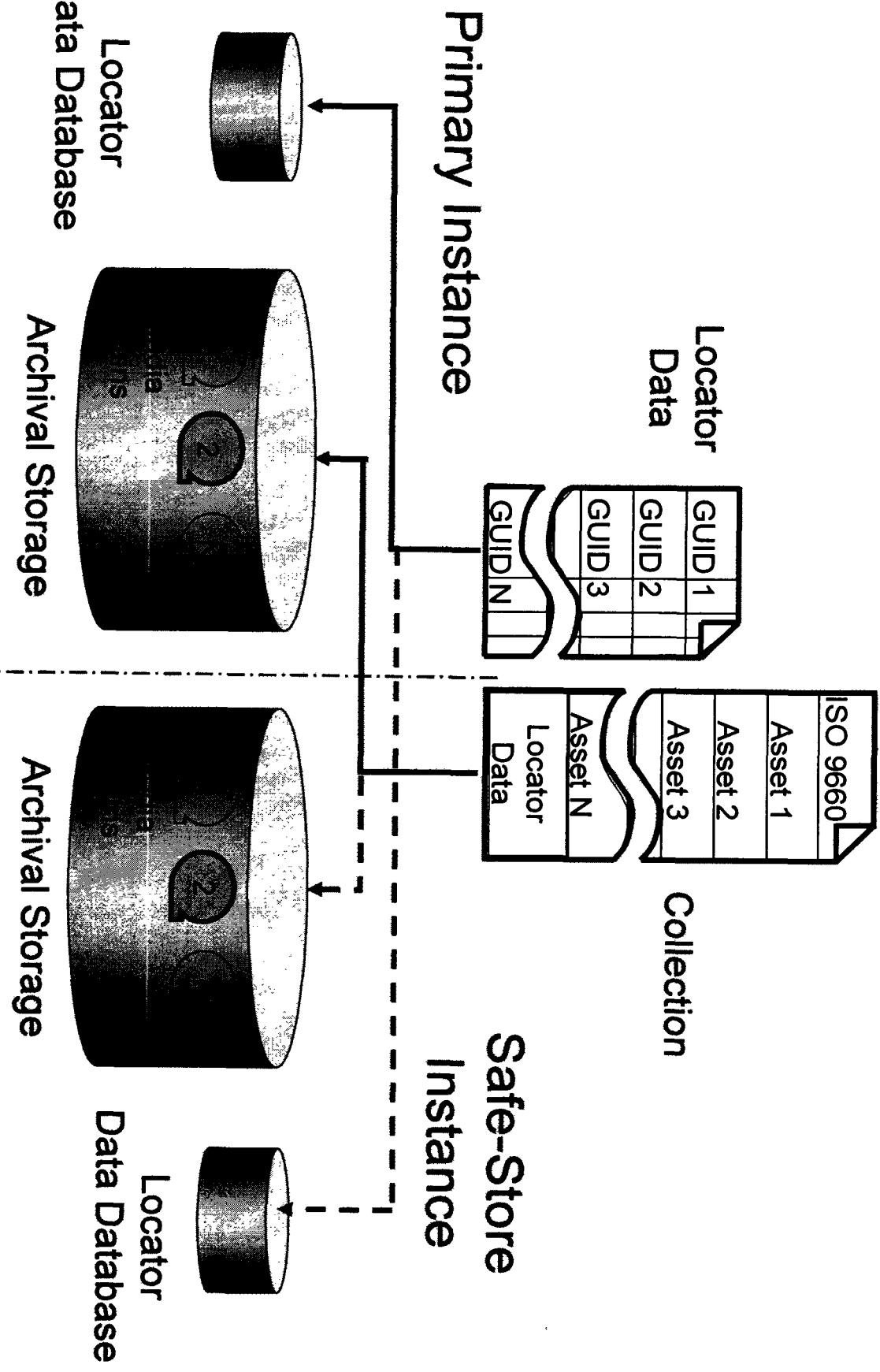


Archival Storage Design Details: Virtual File System Concepts



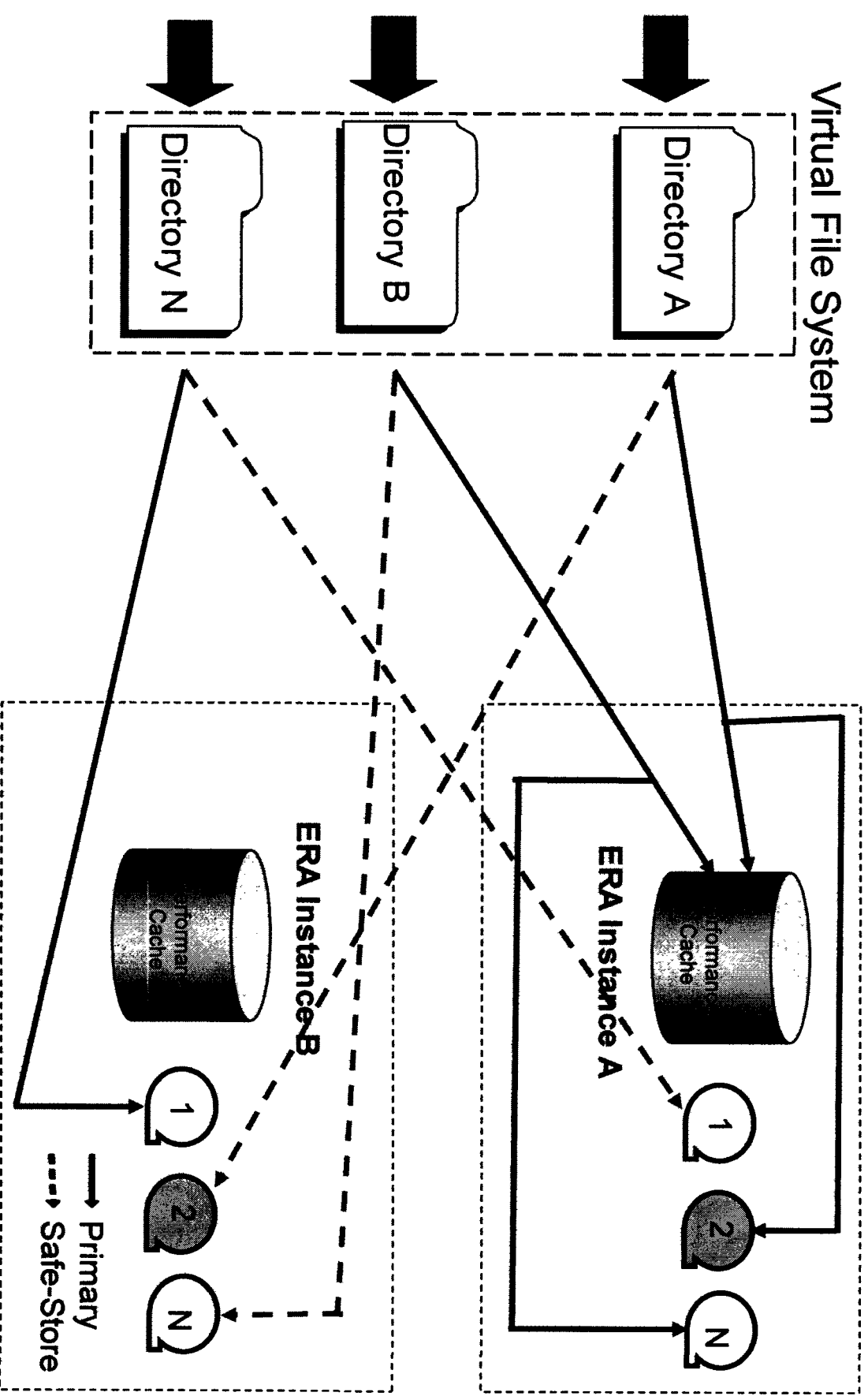
Archival Storage Design Details

Record Safe-Store Concept (continued)



Archival Storage Design Details

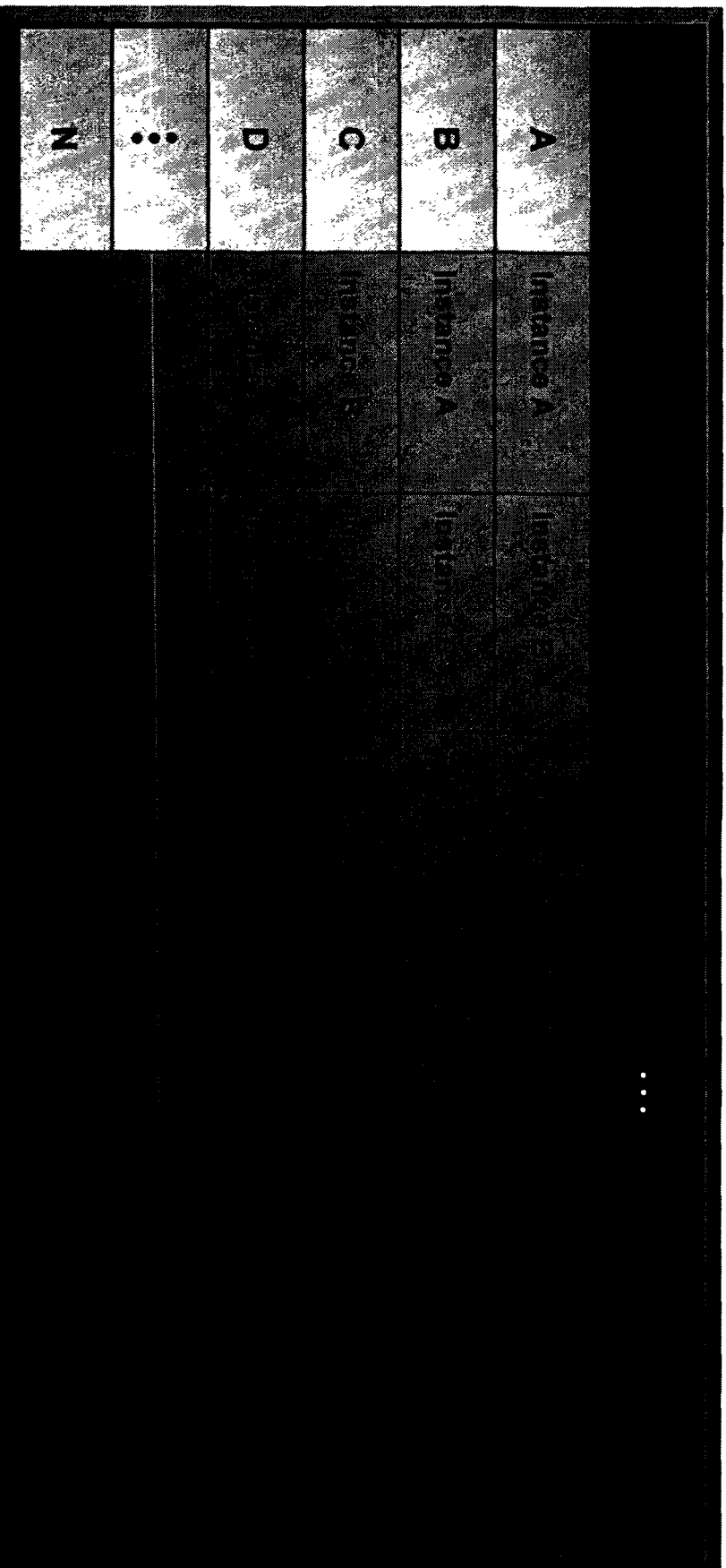
Record Safe-Store Concept



Archival Storage Design Details

Proposed Directory Structure

Scalable to virtually unlimited number of directories
Redundancy of records (i.e. Primary & Safe-Store)
Policy based Levels of Service
Policy based data segregation (i.e. Title 13)

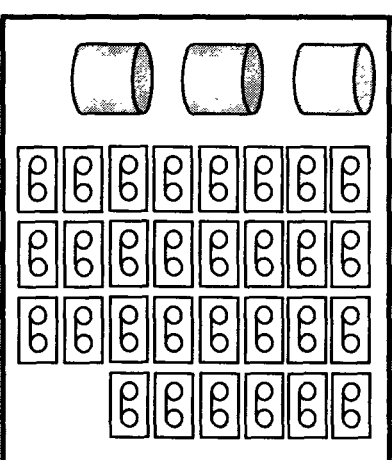
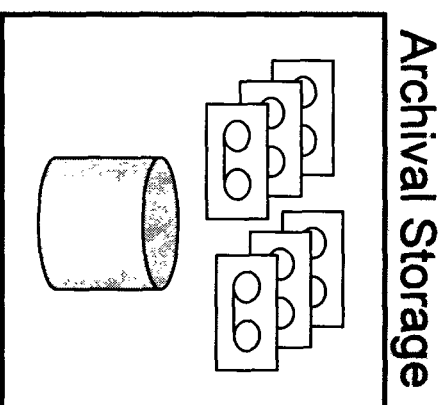
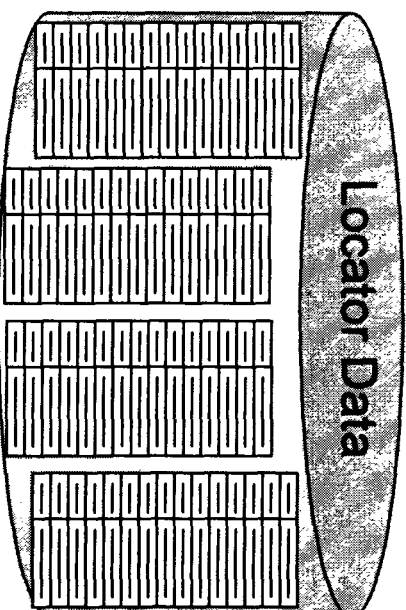
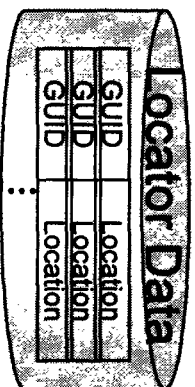


Archival Storage Design Details

Scalability Characteristics

Locator Data

- Contained in Embedded Relational Databases
- Maps GUID to Physical File Location
- 1 Database per Directory



Scales Incrementally as required

Scales to

- 255 Billion Entries per Directory
- 100 Directories per Storage Management Server
- Unlimited Number of Servers
- 40 Directories = 10 Tera objects

Archival Storage Design Trades

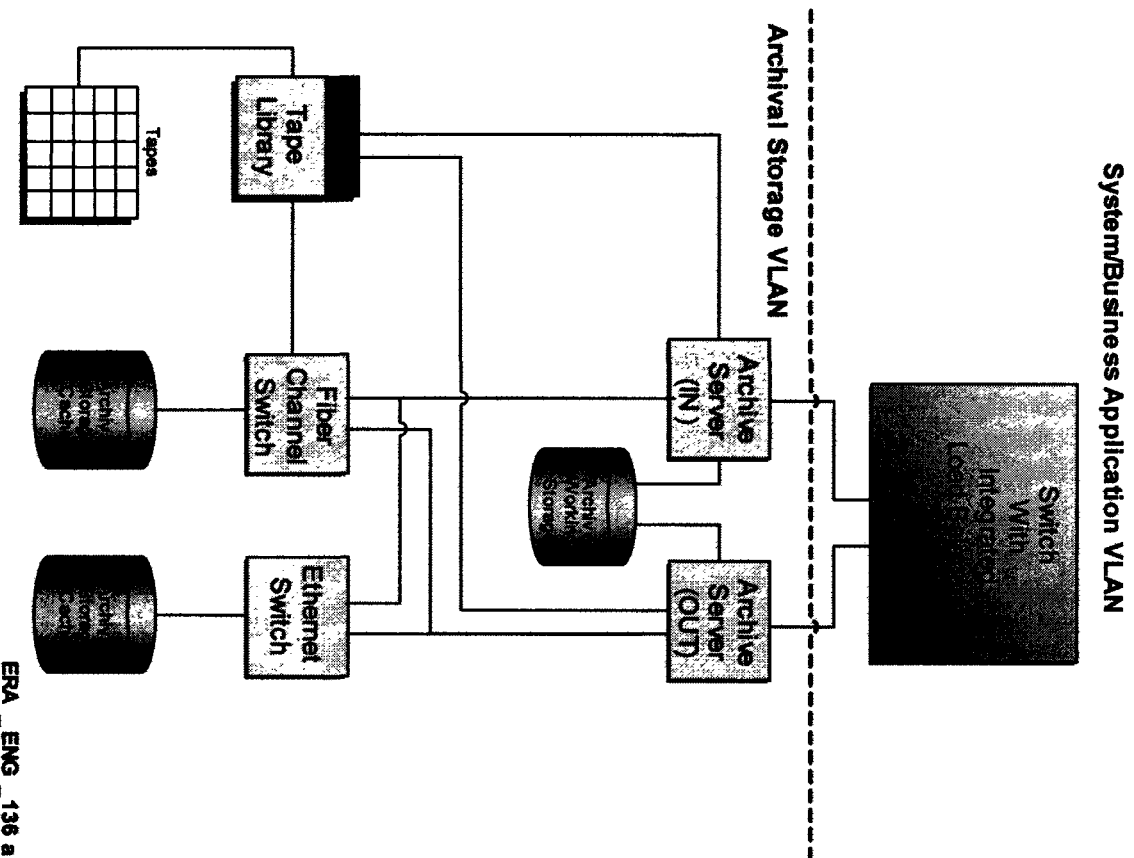
Conducted cost performance and suitability study

Considered trade-offs between the various options

Reduced the storage technology options to those suitable for a long term archive

Selected an architecture that meets the price performance needs of NARA

Archival Storage Physical Design

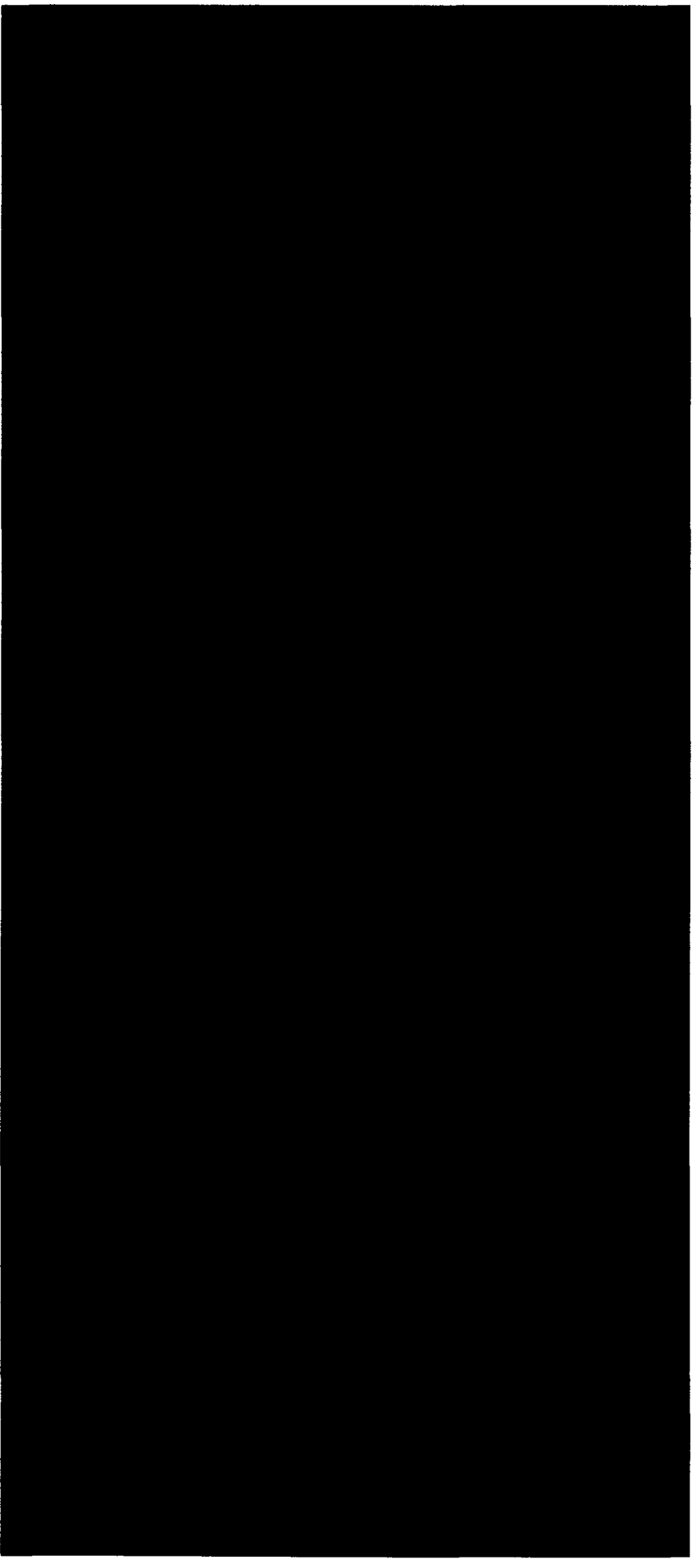


Archival Storage Architecture

- Selected a hybrid storage approach using a combinations of storage technologies
- Switched fabric VLAN allowing high bandwidth access to and from many locations
- Network attached RAID for performance-based caches
- Magnetic tape being used for long term archival storage

ERA _ ENG _ 136 a

Archival Storage VLAN COTS Products



Archival Storage – Conclusions

Active Safe-Store

- Provides archive backup copies of its assets, continuously preparing to recover from hardware, software, network, and Instance failures

Virtual File System

- Provides a consistent interface to the other services.
- Request to read an asset only needs to know the globally unique identifier, GUID, of the asset and not the physical location or media chain of the asset

Abstraction of Mass Storage

- Provides a medium and supplier-independent storage solution
- Allows new device and media technology to be introduced without changes to the other services

Incremental Scalability

- Virtual File System and Abstraction of Mass Storage
- Provide for the incremental addition of storage media
- Addition of ERA instances to ensure a solution that will scale as required

Archival Storage RIDs

RID-LMC00134 COTS Storage Selection

LUNCH

ERA SDR - DAY THREE

Dissemination Design

May 11, 2005

Dissemination Design Agenda

Description of Functionality

Key Requirements

Functional Architecture

Service Design

Physical Design

Initial Product Selections

RID Discussion

Dissemination Description

Dissemination provides

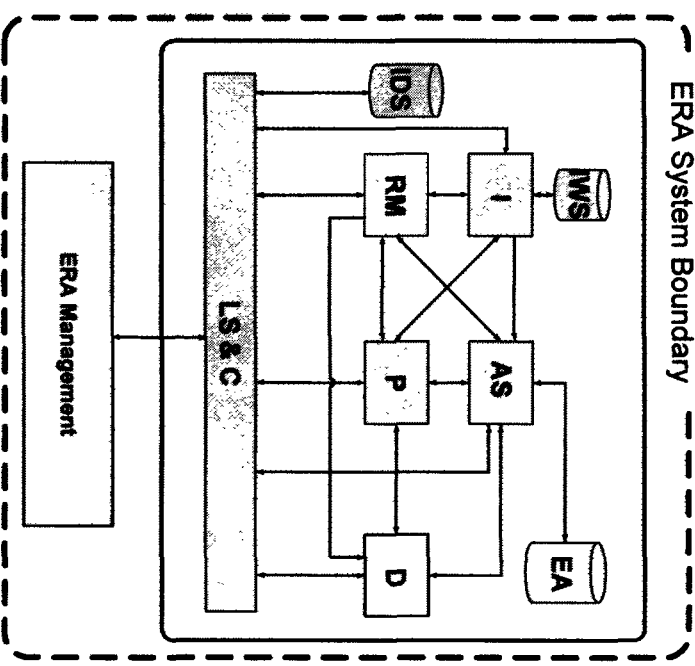
- Functionality to manage search and access requests for assets within the ERA System**
- A framework to enable the use of multiple search engines offering a rich choice of searching capabilities across assets and their contents**
- A framework to enable the use of multiple viewers for preservation of assets**

Dissemination Driving Requirements

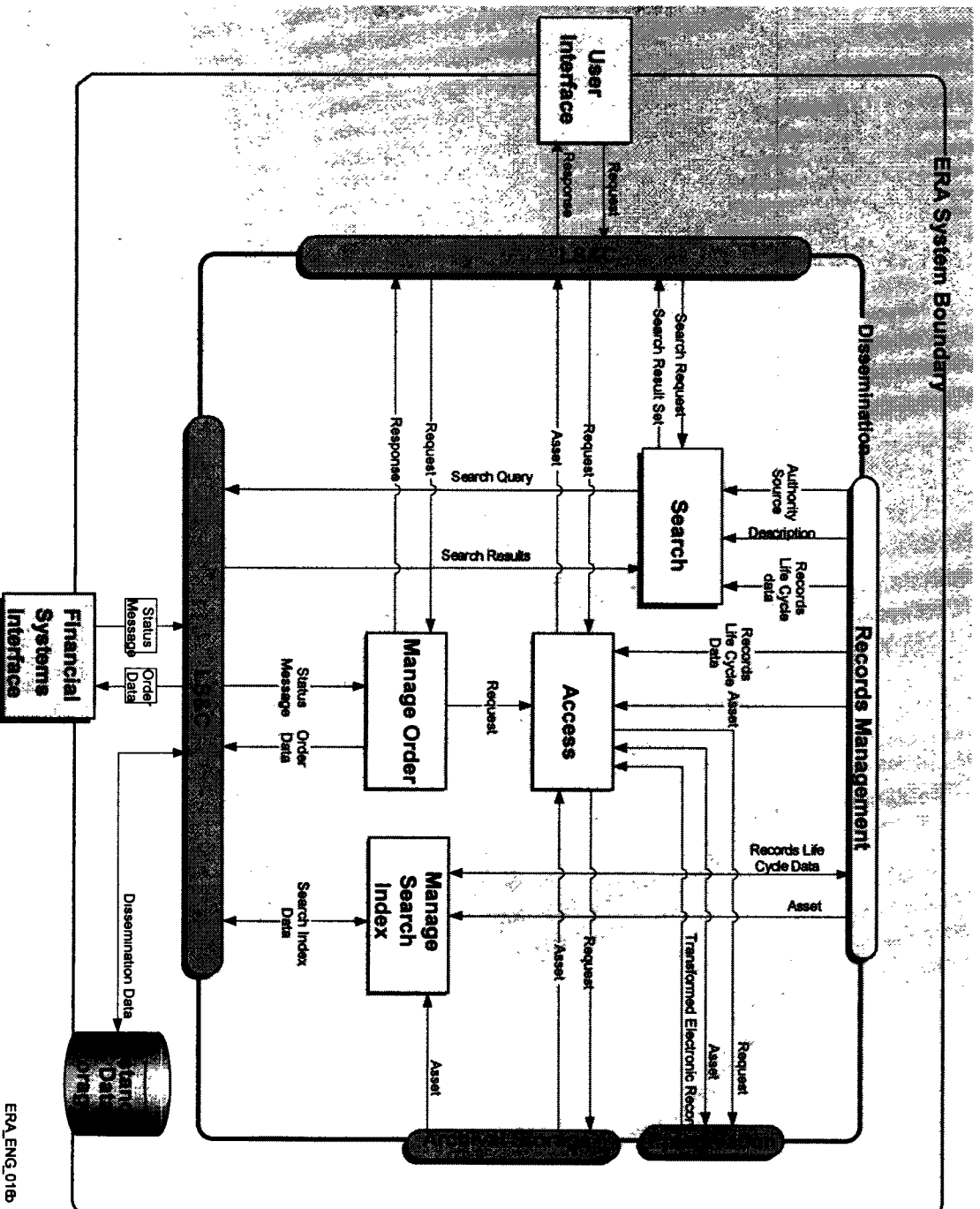
LM19 – Flexible framework to support different kinds of searches

LM20 – Flexible framework for providing access to assets

LM20.4, LM20.5 – Interface between Dissemination and digital adaptation services in Preservation



Dissemination Functional Architecture



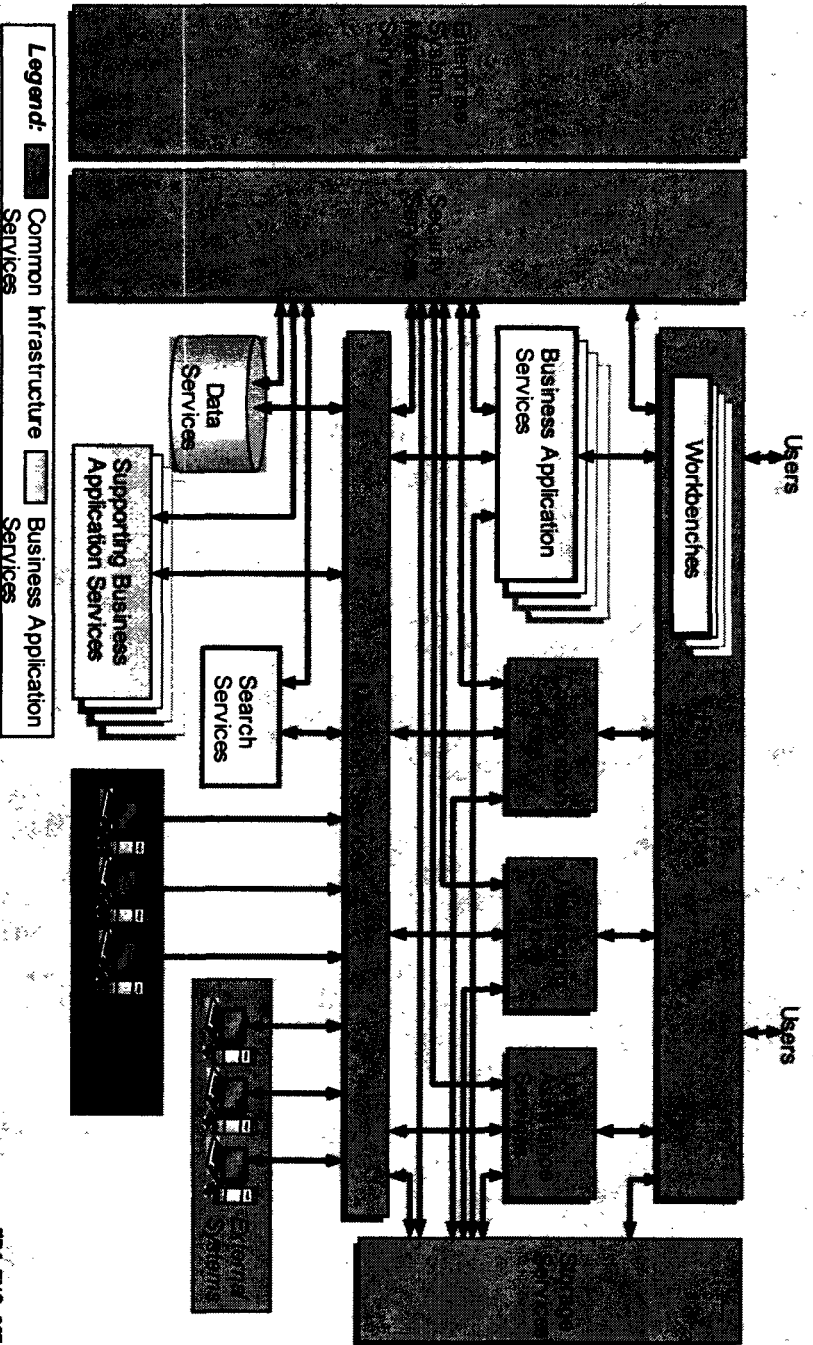
Key Features:

- Search Framework
- Access Framework
- Manage Orders
- Manage Search Indexes

Dissemination within SOA

Dissemination includes

- Workbenches
- Business Application Services
- Search Services



Dissemination Services

Consumer	Provides user with current and historical information on searches, including information on the status of the search, the results of the search, and the location of the search results.
Access	
Manage Orders	
Search	

Dissemination Design Highlights

Search Framework

- Supports different kinds of searches, including searching by arrangement, description content, record life cycle data, record content

Search Strategy

- Hybrid approach which offers different trade-offs between comprehensiveness, responsiveness, and cost

Access Framework

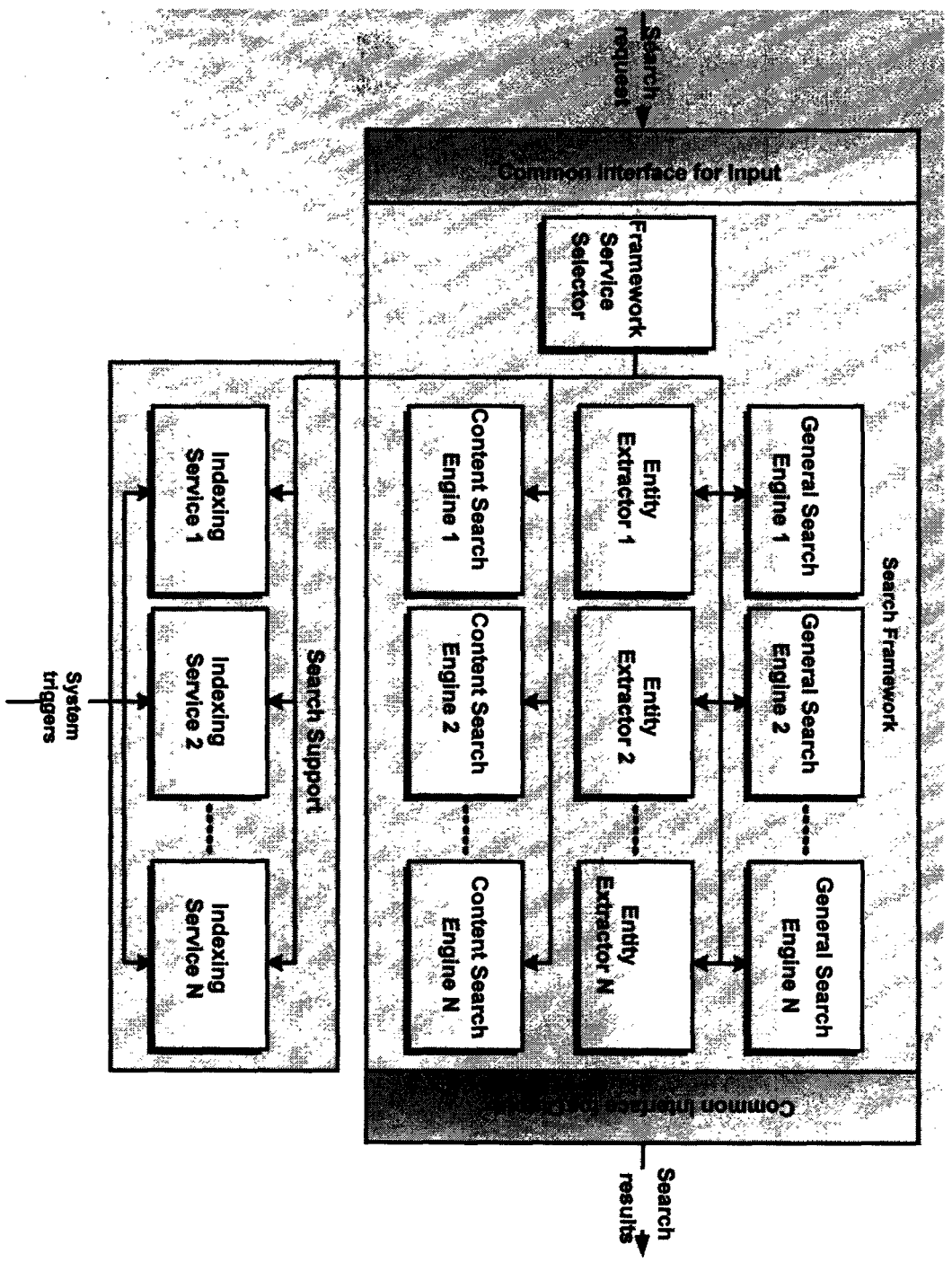
- Supports output and presentation of different kinds of assets

Certification

- Provides a way to certify accessed assets

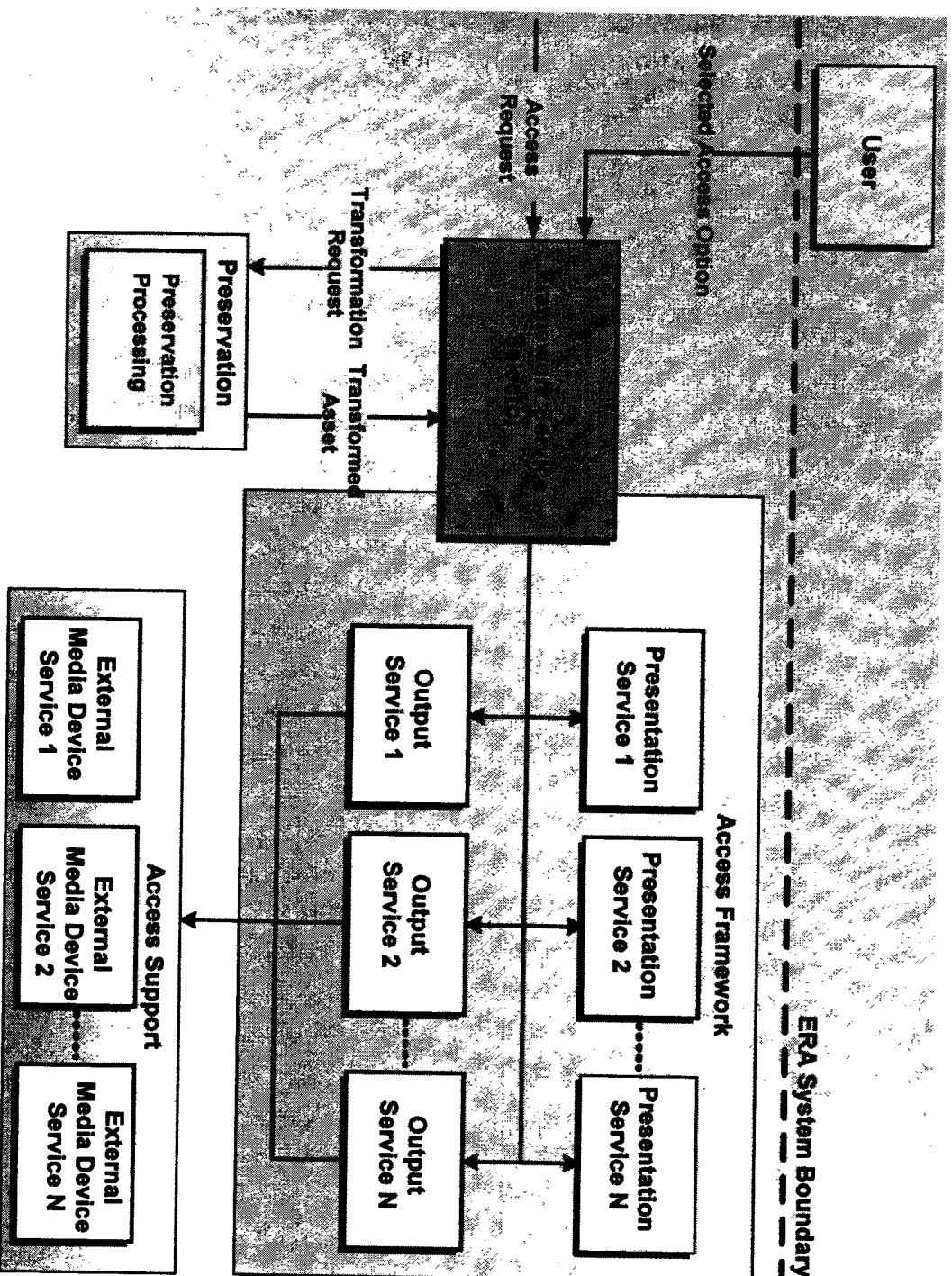
Dissemination Design Details

Search Framework



- Key Features:**
- Multiple Search Engines
 - Multiple Entity Extractors
 - Multiple Indexing Services
 - Select “best-fit” search for particular needs

Dissemination Design Details Access Framework



Key Features:

- Multiple Presentation Services
- Multiple Output Services
- Multiple Media Devices
- Integrated Preservation Processing
- Select “best-fit” access for particular needs

Dissemination Design Trades

Use COTS search engine products to facilitate search

Use a framework to abstract ERA from details of specific COTS search engines

Use COTS targeted entity extraction to extract metadata from content automatically

Investigated various search approaches

- Content**
- Concept**
- Hierarchical**
- Record life-cycle data**
- Data types**

Dissemination Design Models

Established initial models and framework for Search

Future modeling work will refine the models and provide inputs such as...

- What are the storage requirements for the proposed search strategy?**
- How will the system meet the dissemination requirements?**
- How many searches will the system have to process in a given year?**
- What are the server (CPU, memory) requirements for Search?**
- How many dissemination requests will the system have to process in a given year?**

Dissemination Physical Design

Dissemination services are implemented on the System/Business Applications VLAN, which is described in the Local Services & Control Design charts

Dissemination – Conclusions

Provide Search Framework

- Supports different kinds of searches
- Encapsulates proprietary details of COTS products
- Allows new search engines to be added over time

Optimal Search Strategy

- Balances trade-offs between comprehensiveness, responsiveness, and cost
- Offers a reasonable approach to providing content search

Provide Access Framework

- Supports different kinds of access viewers
- Encapsulates proprietary details of COTS products
- Allows new access viewers to be added over time

Dissemination RIDs

RID-LMC00121 Search Capabilities

RID-LMC00122 Mediated Searches

RID-LMC00124 Access of Non-Electronic Records

RID-LMC00126 Order Data in the IRD

BREAK

ERA SDR -- DAY THREE

ERA Management



May 11, 2005

ERA Management Description

ERA Management provides

- Centralized system monitoring services**
- Centralized services to manage the ERA System**
- Centralized configuration management services**
- Centralized ERA Help Desk**

ERA Management Design Agenda

Description of Functionality

Key Requirements

Functional Architecture

Service Design

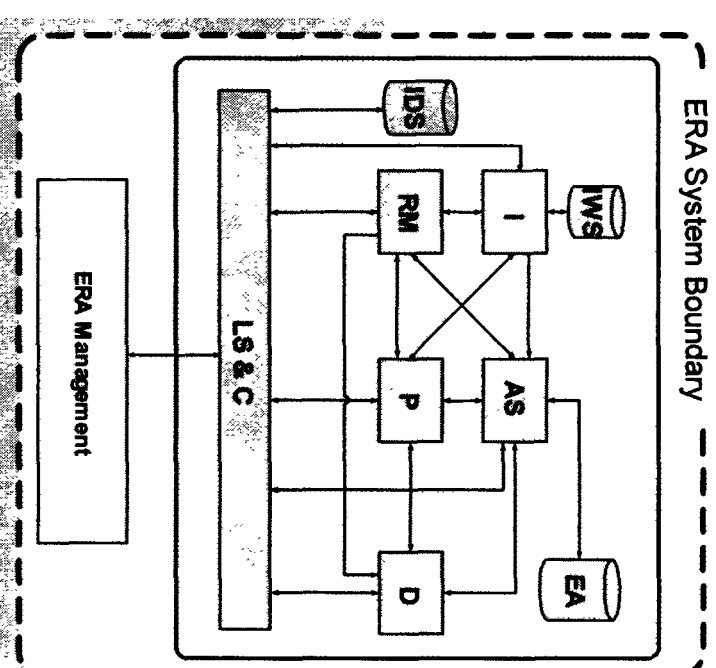
Physical Design

Initial Product Selections

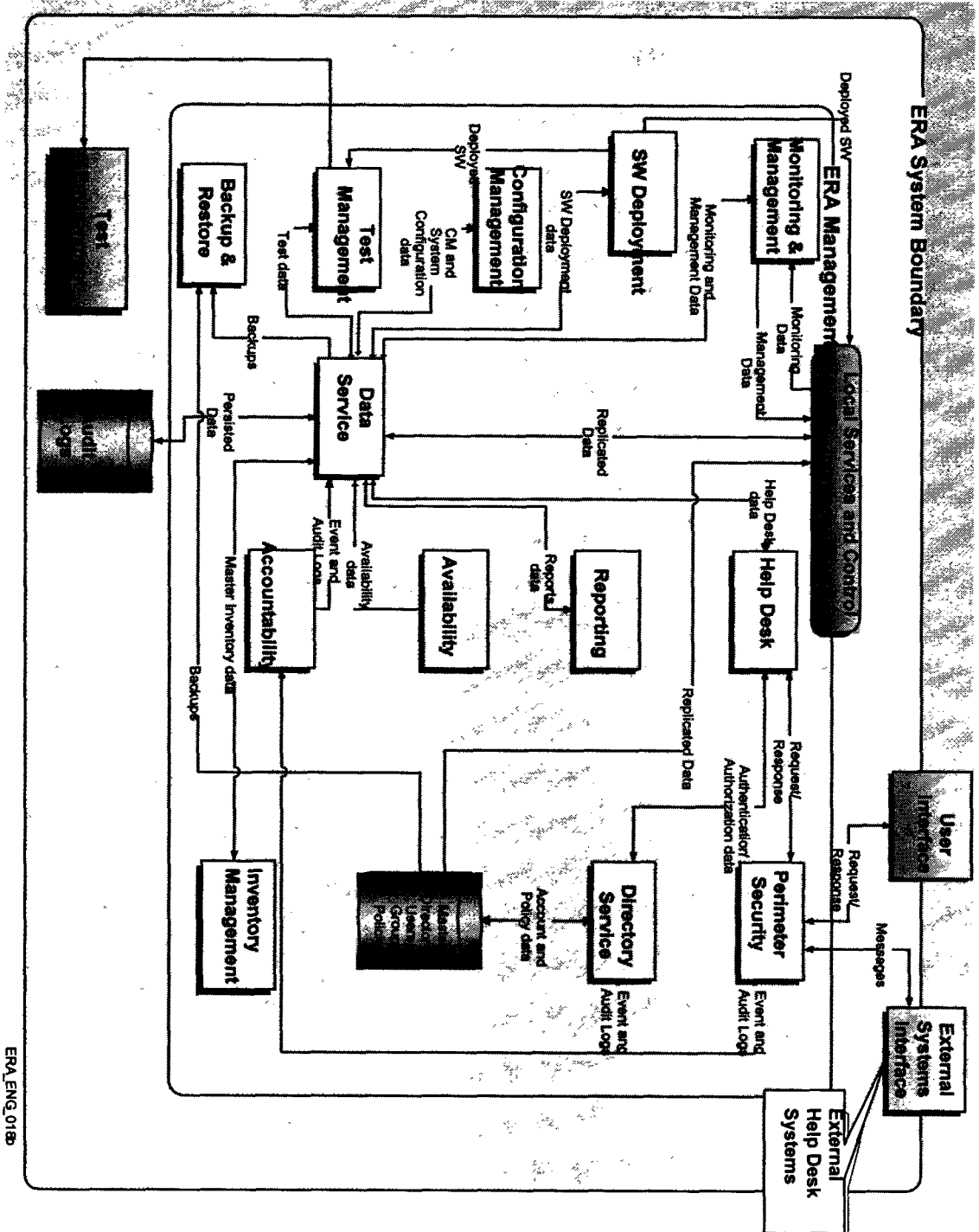
RID Discussion

ERA Management Driving Requirements

- LM22 – Centralized Directory Service**
- LM23.1 – Integrated user assistance**
- LM23.3 – Interface with the NARA Help Desk**
- LM25 – Centralized roll-up of event logs**
- LM26 – Comprehensive reporting**
- LM27 – Centralized monitoring, management, and remote administration**
- LM28 – Centralized tracking of all the ERA System assets**
- LM32 – Clustering of servers for each Instance and the global load balancing among Instances**



ERA Management Functional Architecture



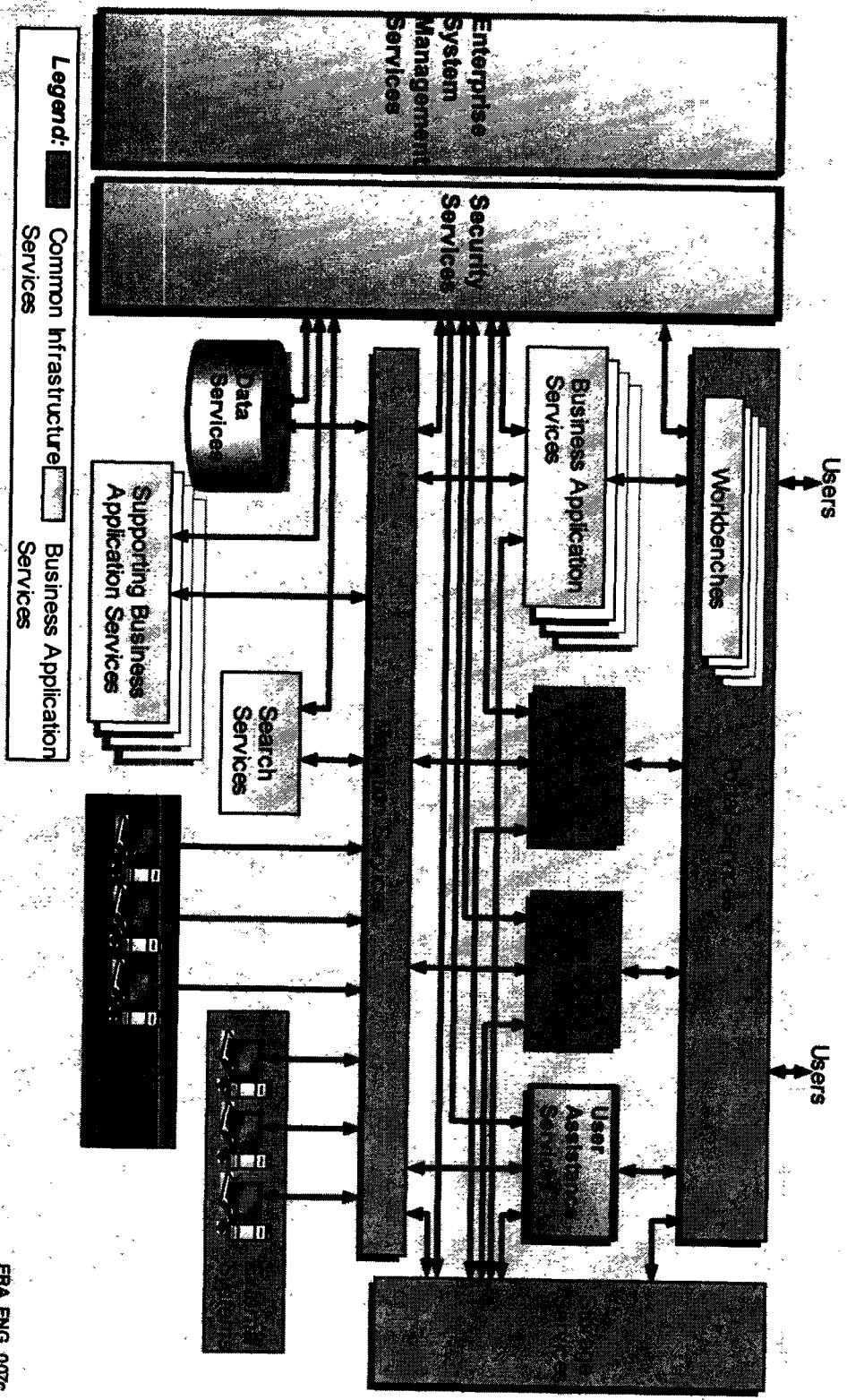
ERA_ENG_0180

- Key Features:**
- Directory Master
 - Network and System Monitoring
 - Software Deployment
 - Inventory Management
 - Help Desk Reporting

ERA Management within SOA

ERA Management includes

– Centralized Common Infrastructure Services



ERA_ENG_007c

ERA Management Services

Perimeter Defense	Provides help with configuring and maintaining intrusion detection and prevention systems.
Directory Service	Provides help with configuring and maintaining directory services.
Accountability	Provides help with configuring and maintaining logging and auditing systems.
Data Service	Provides help with configuring and maintaining data services.
Monitor and Manage Sensors	Provides help with configuring and maintaining sensors.
Monitor and Manage Alerts	Provides help with configuring and managing alerts.
Monitor and Manage Events	Provides help with configuring and managing events.

ERA Management Services

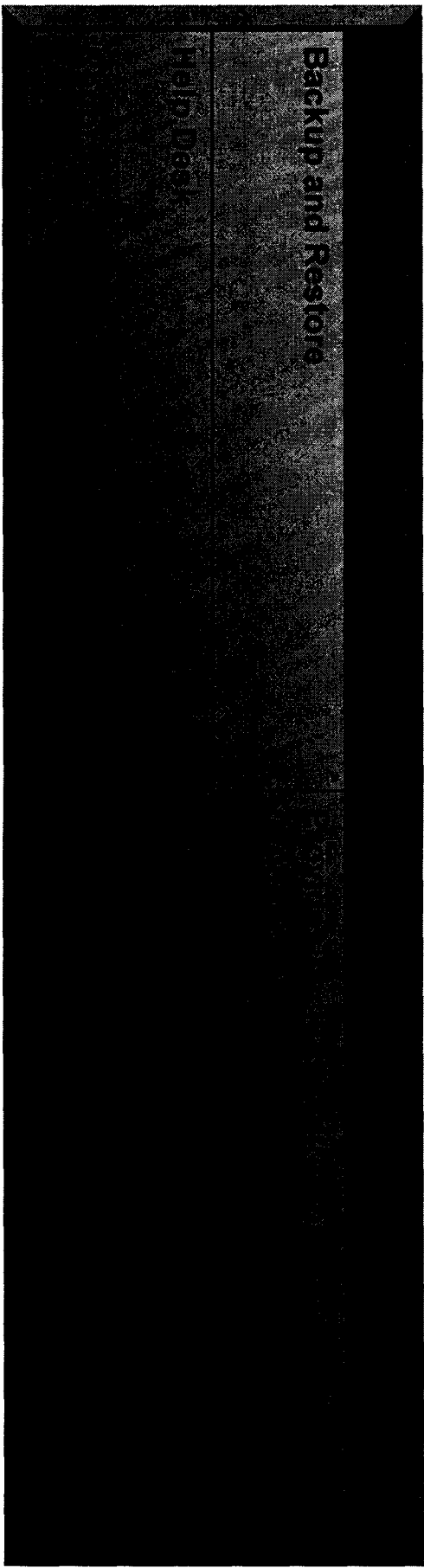


The page content is almost entirely obscured by a large black redaction box. Only a few faint labels are visible on the left side of the page, which appear to be part of a table or list structure. These labels include:

- Availability
- Configuration Management
- Test
- Software Deployment
- Management

On the right side of the redacted area, there is some very faint, illegible text that appears to be a header or a note, possibly starting with "Provide support for...".

ERA Management Services



ERA Management Design Highlights

Accountability

- Includes error, event, and audit logs, and tools for interactive query and analysis
- Centralized tools pull together the logs from the individual operating systems, COTS products, and security appliances into a cohesive and comprehensive view

System Management

- Provides centralized monitoring and management for servers, applications and databases, storage, and networks
- Provides a centralized monitoring and management of facility and network loads, and the balancing of user and interface load across sites
- Provides automated deployment of configuration managed hardware and software

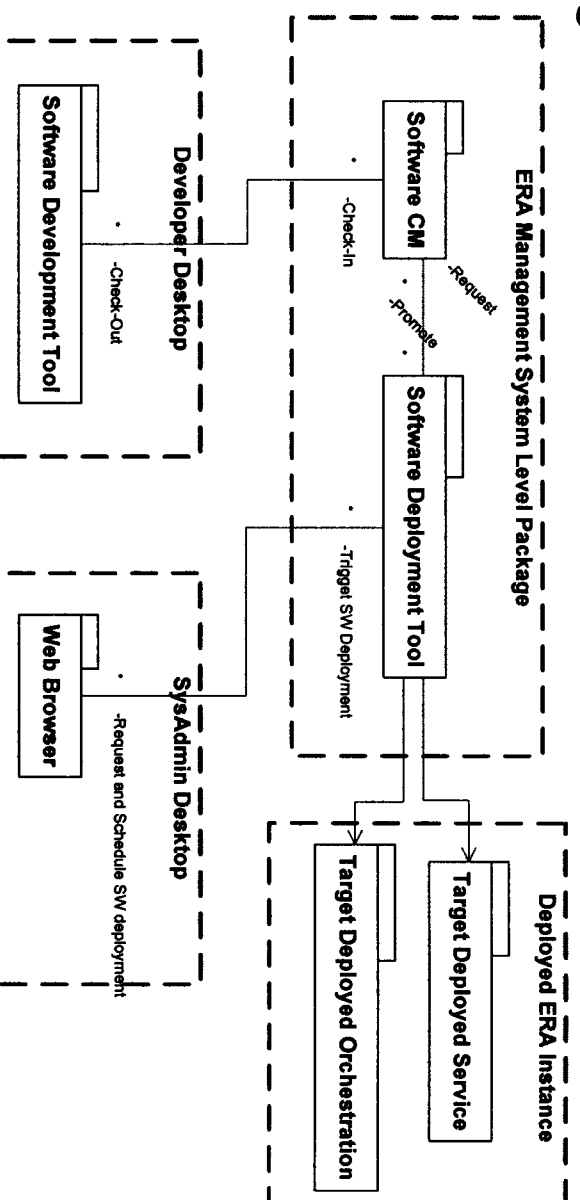
Test Management

- Provides tools for the management of test data, test scripts, test execution and problem tracking in the test environment

ERA Management – System Management

System Monitoring and Management tools and processes

- CA Unicenter
 - Cisco Works
 - StorageTek Global Storage Manager
 - Leverage processes and configurations from existing datacenters
- ## Software Deployment tools and process
- Deployment on service-by-service basis
 - CM controlled process



ERA Management Design Highlights

Inventory Management

- Media containing electronic records and assets are tracked and audited
- Accountable inventory are tracked and cost accounted
- Consumable inventory is not tracked after receipt

Help Desk

- Provides user assistance and problem reporting for the ERA System
- Includes a continually-updated Knowledge Base
- Includes a “warm handoff” interface to the NARA Help Desk for internal NARA IT assistance and for referring researcher assistance requests to the archival staff

Reporting

- Includes the capability for standard (canned) reports and ad hoc reports

ERA Management Design Trades

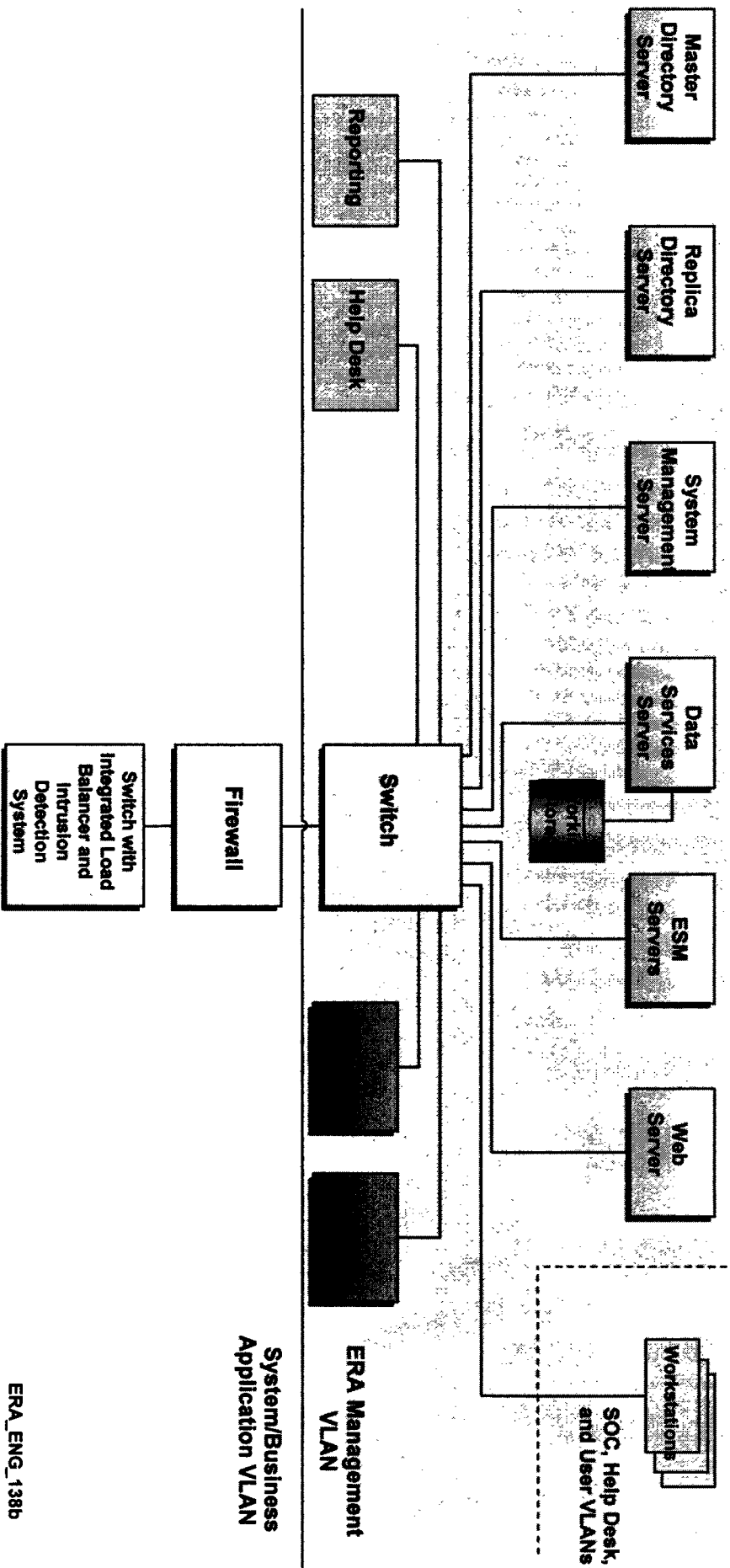
Implement a Service Oriented Architecture

- Collected common infrastructure services into core packages
- LS&C provides distributed core services that are leveraged by business service components
- ERA Management provides centralized core services to manage business service components

Leverage COTS products to provide core infrastructure services

- Enterprise Systems Management
 - Servers, applications, databases, storage, network
 - “Manager of Managers” to provide:
 - Event correlation
 - Unified dashboard with consistent look and feel
- Data replication to enable centralized management and distributed operations

ERA Management Physical Design



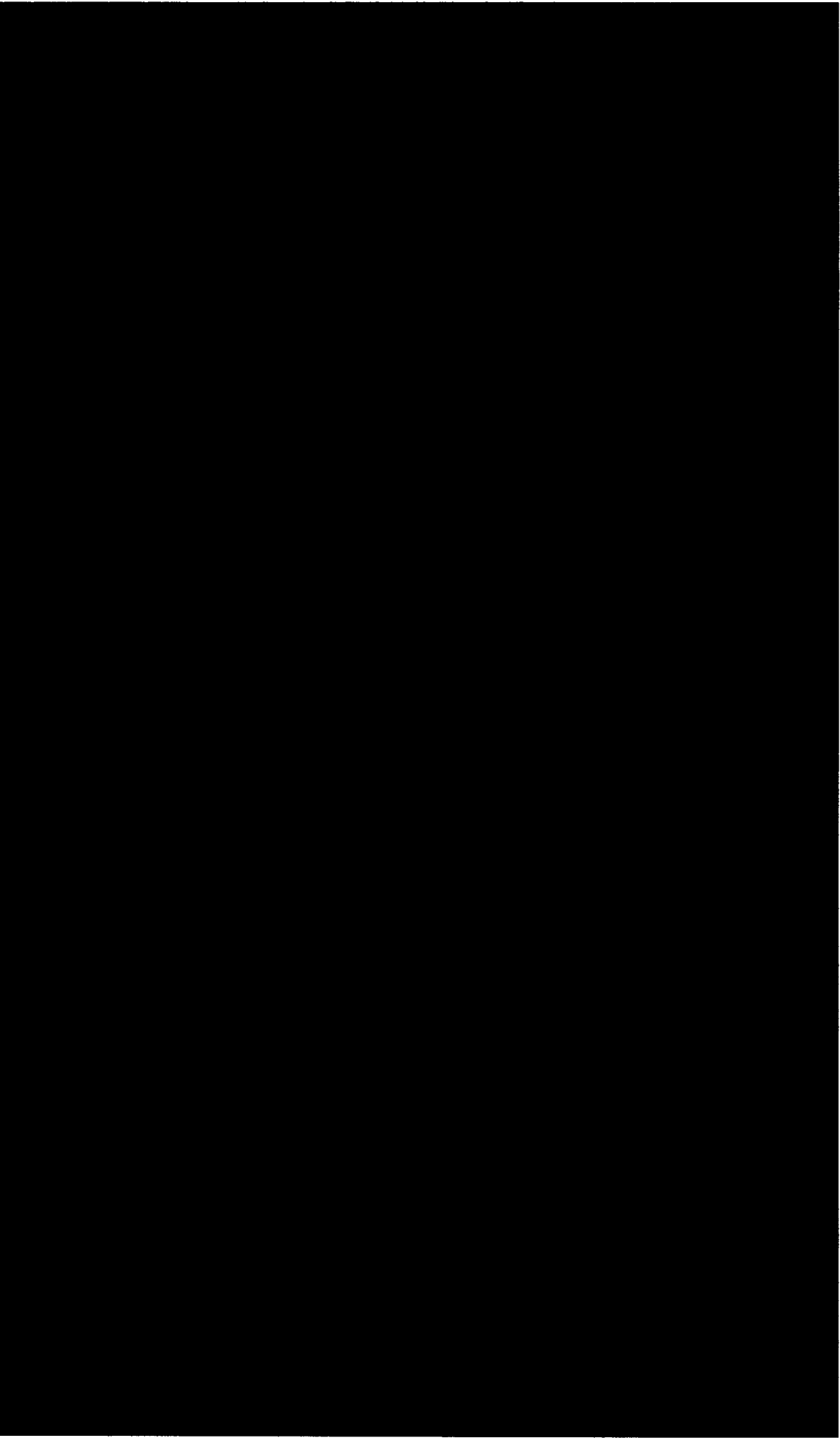
ERA_ENG_138b

ERA Management Physical Design

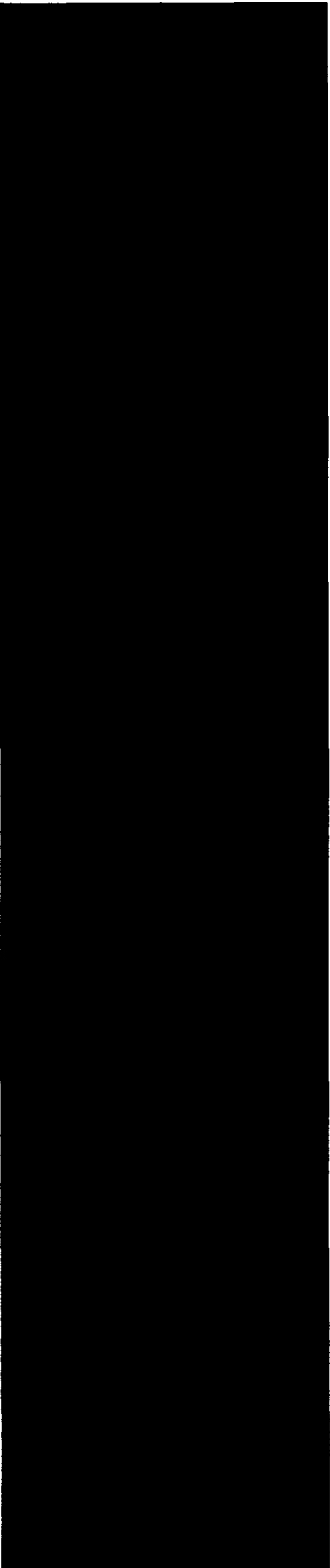
ERA Management VLAN

- Exists at one primary and one fail-over facility within each Federation**
- Contains all of the monitoring and management infrastructure service servers required to remotely monitor and manage the Instances within a Federation**
- Includes partitioned servers**
 - Allows unrelated service components to be hosted on the same physical server in a manner that is convenient for deployment, scalability, operations, and management**
 - Includes dedicated servers where needed for performance or clustering**
 - Scales vertically and horizontally to larger or smaller configurations**
- Connected to the rest of the ERA System with an out-of-band network, and is isolated with a firewall**
- Includes administrative workstations on fire-walled LAN segment**

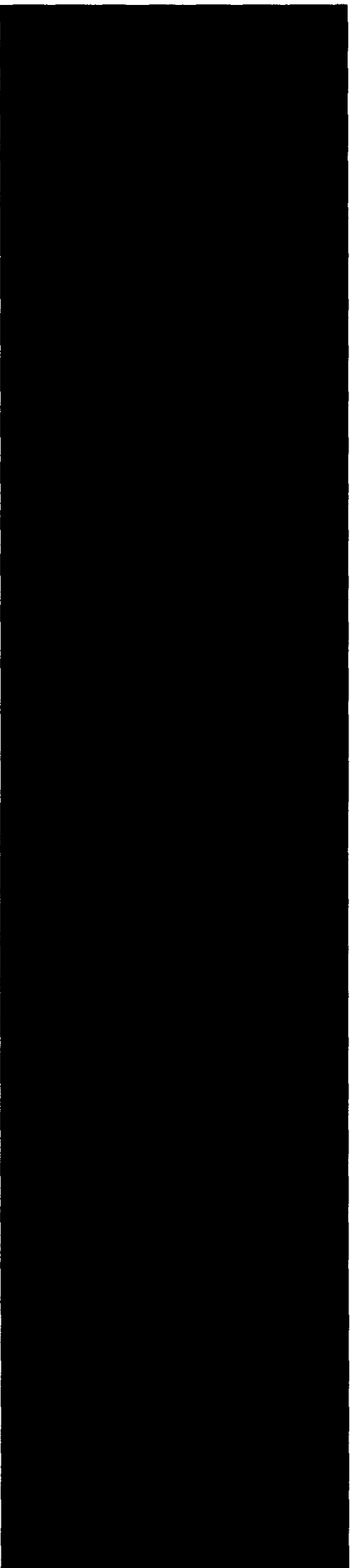
ERA Management VLAN COTS Products



ERA Management VLAN COTS Products



ERA Management VLAN COTS Products



ERA Management – Conclusions

Centralized System Management

- Provides a unified view into the health and status of a federation**
- Ensures operational performance levels are maintained**
- Tracks problems and their resolution**
- Provides automated software deployment**

Optimal Systems Management solution that balances automation, centralized management, and cost

ERA Management RIDs

RID-LMC00132 Use of MS2003 Products

ERA SDR - DAY THREE

External Interface Design



May 11, 2005

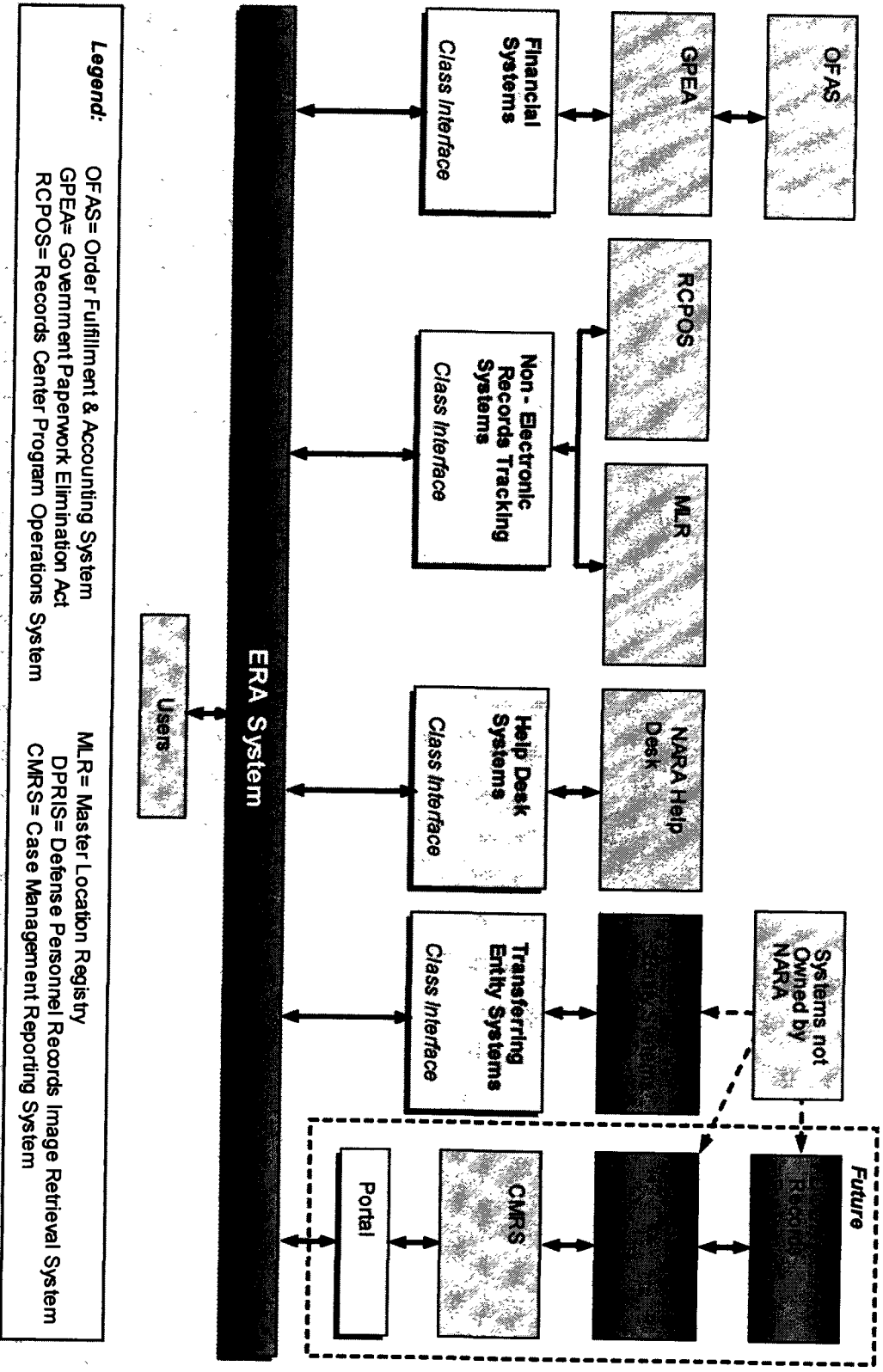
Interface Control Documents

Establish standards, protocols and messages for exchanging information between the Interface Class and the ERA System

ICD for each interface class:

- Financial Systems**
- Non-Electronic Records Tracking Systems**
- Help Desk Systems**
- Transferring Entities Systems**

ERA System External Interfaces



ERA_ENG_0036

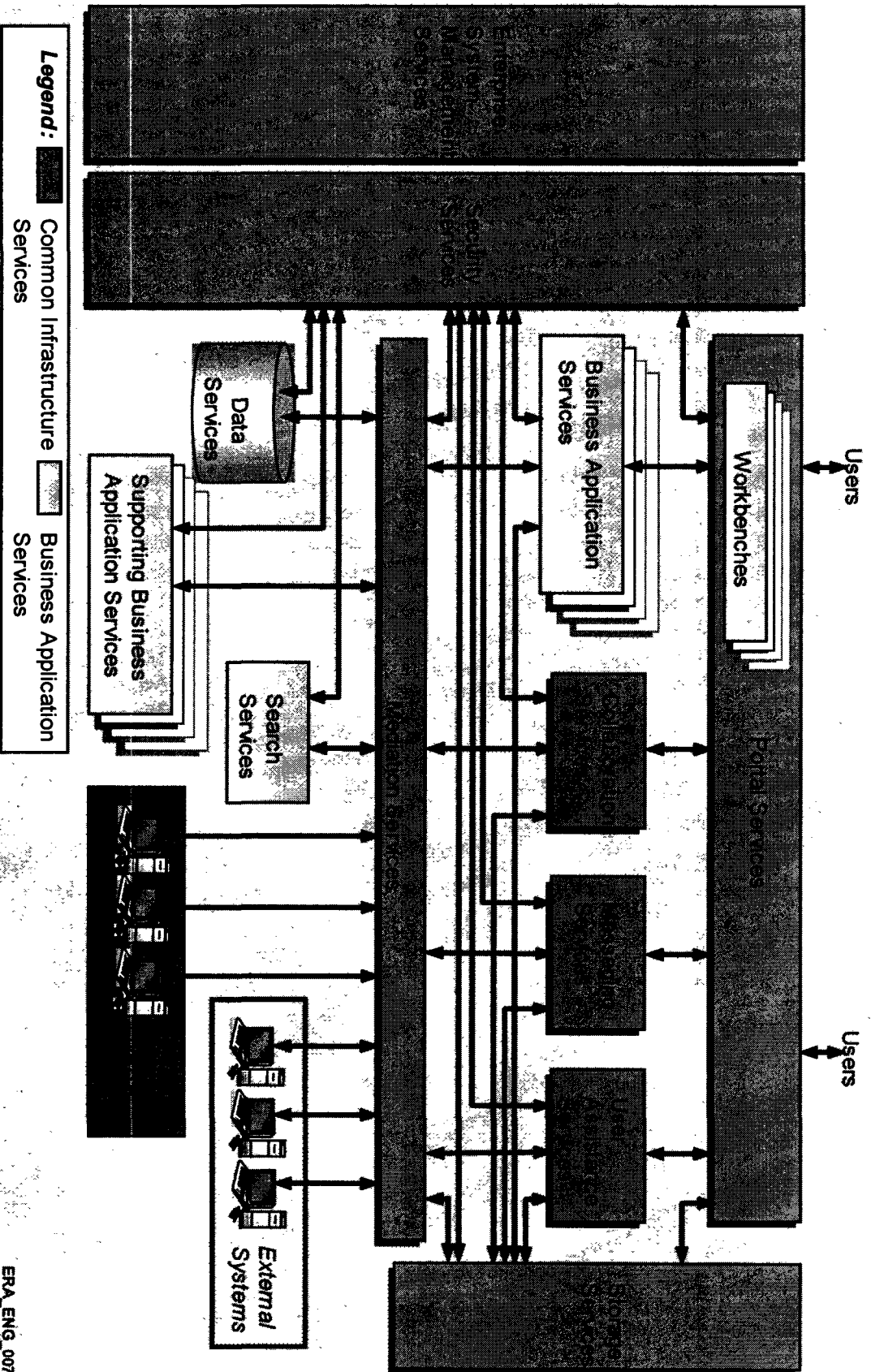
External Interfaces within SOA

ERA System provides service-based interface to external systems

ERA Mediation Services support interface connections:

- Synchronous**
 - Near real time interaction required
 - Implemented via request/response XML/SOAP web service
 - Calling system “blocks and waits”
 - Error handling includes timeout
- Asynchronous**
 - Near real time not required, but excessive response wait time not tolerable
 - Implemented via XML/SOAP web service

External Interfaces within SOA



ERA_ENG_007c

Financial Systems Interface: Messages

Asynchronous – Send:

- Product Order Delivery Information**
- Product Order Shipping Information**
- User Identification**
- Updated User Identification**
- Resources Required**
- Order Produced**
- Order Cancelled**

Asynchronous – Receive:

- Total Cost of Order**
- User Information**
- Authorization to Proceed**
- Denial of Authorization**

Financial Systems Interface: Messages (continued)

Synchronous – Send:

- Service Order Billing Information**
- Service Order Payment Information**
- Service Order Delivery Information**
- Service Order Shipping Information**
- Notification**

Synchronous – Receive:

- Notification**

Non-Electronic Records Tracking System Interface: Messages

Asynchronous – Send

- Notification of Schedule Arrival
- Change in Disposition
- Message Receipt Notification

Asynchronous – Receive:

- Notification of Arrival Status
- Returned to Agency
- Scheduled for Destruction
- Records Destroyed
- Message Receipt Notification

Help Desk System Interface: Messages

Asynchronous – Send & Receive

- Help Desk Ticket**
- Help Desk Ticket Resolution**
- Help Desk Ticket Updates**
- Notification**

Transferring Entities System Interface: Messages

Asynchronous – Send

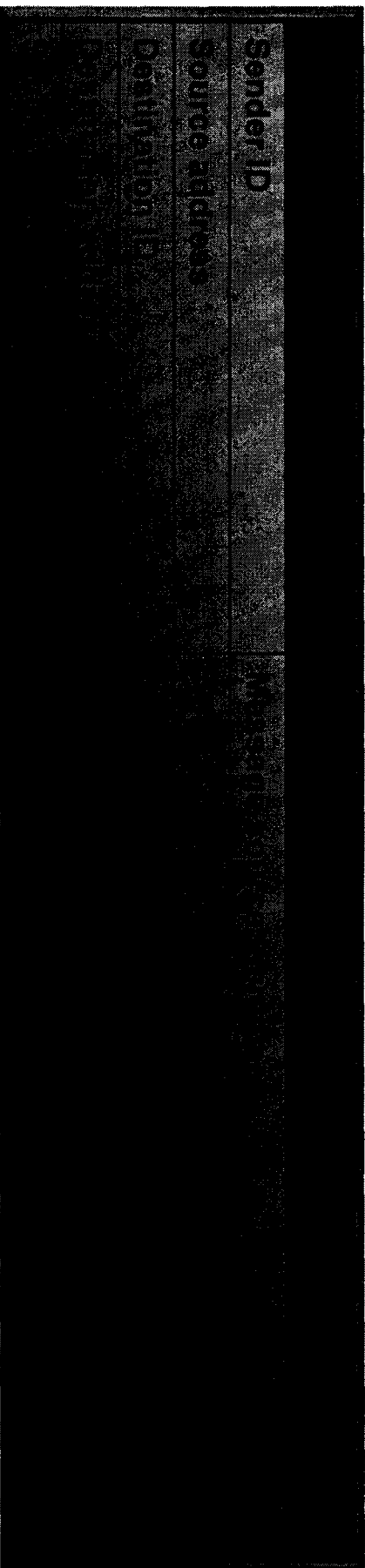
- Records Transfer**
- Description**
- Notification**

Asynchronous – Receive

- Records Transfer**
- Description**
- Notification**

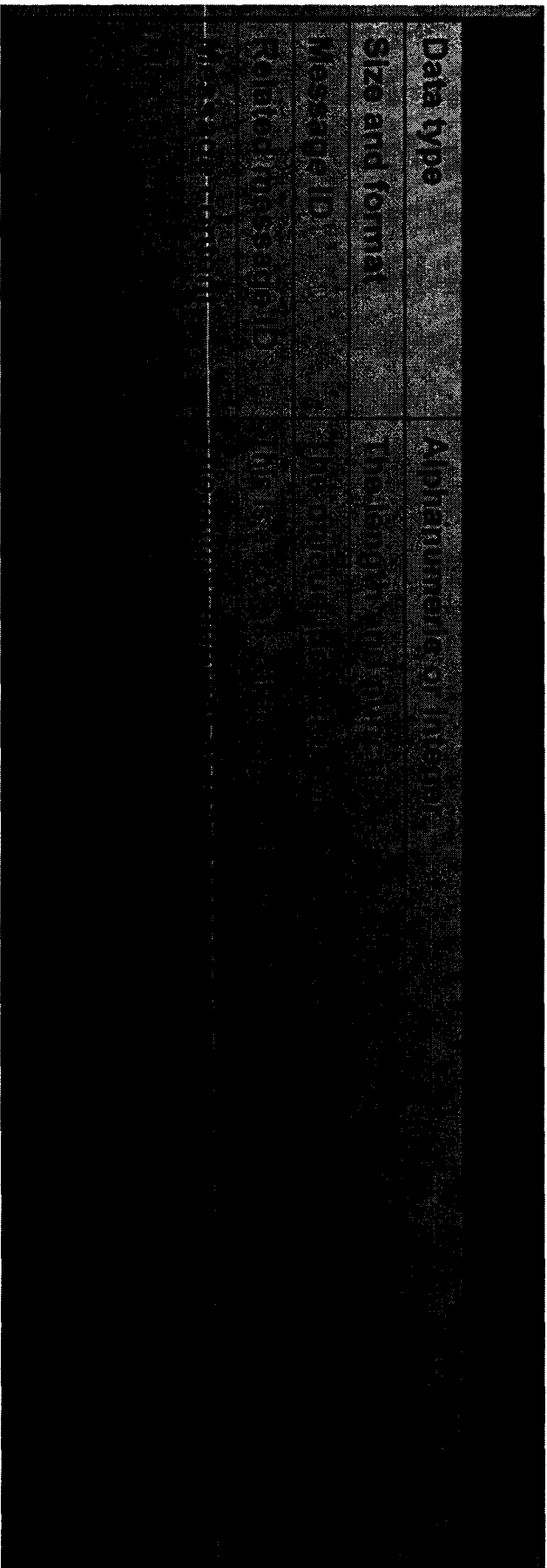
Synchronous Message Schema

Header:



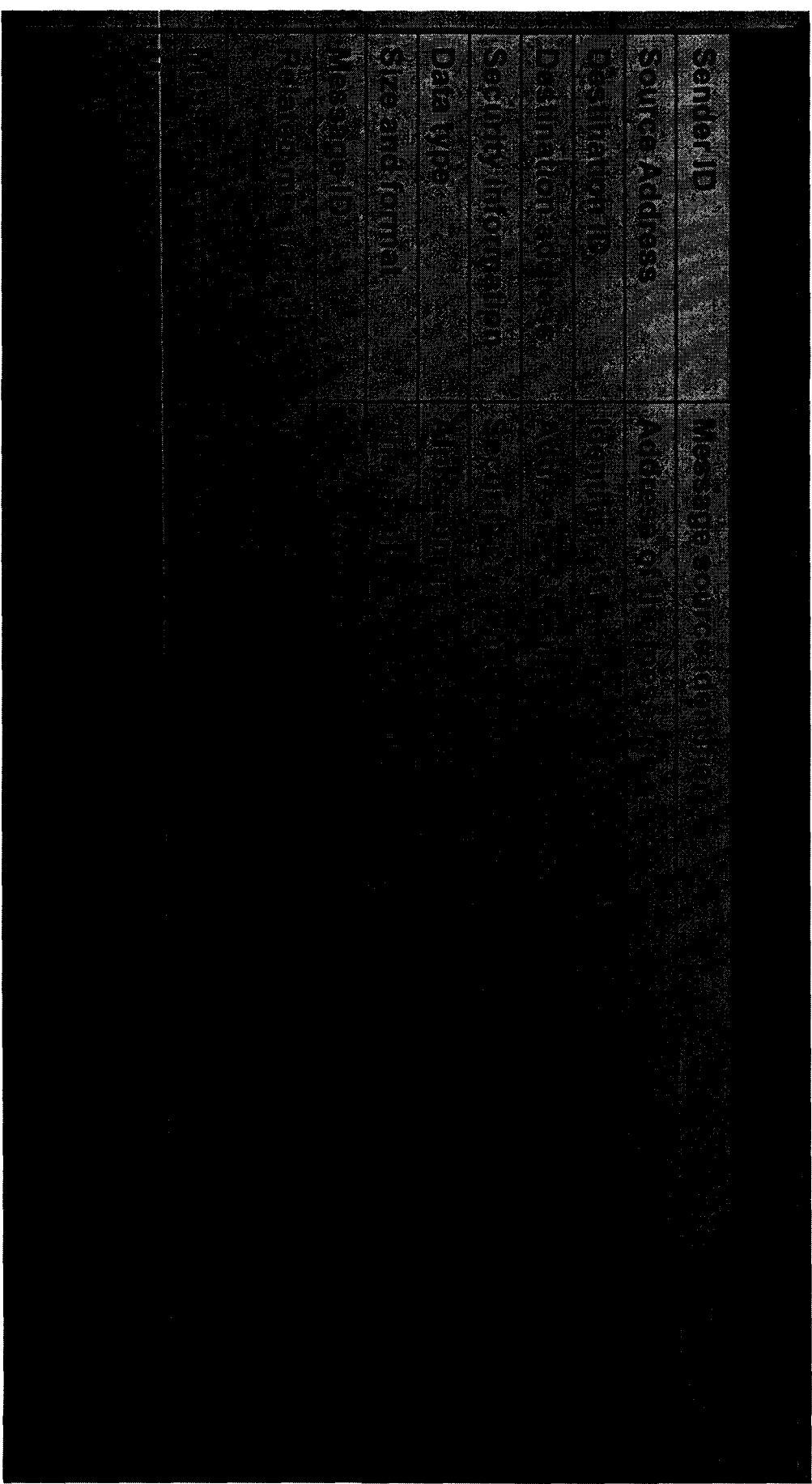
Message:

Data type	Alphanumeric or Integer
Size and format	The length of the message
Message ID	The unique identifier of the message
Related message ID	The unique identifier of the related message



Asynchronous Message Schema

Message:



Sender ID	Message subject identifier
Source Address	Address of the message
Destination ID	Identifier of the destination
Destination address	Address of the destination
Security information	Security information
Data type	Message data type
Size and format	Message size and format
Message ID	Message identifier
Relative priority	Relative priority
Message source	Message source

END DAY THREE