

EMC Solution for eDiscovery Collection



Automate Collect and Hold of Unstructured Content

Challenges of eDiscovery: Are You Prepared?

The U.S. Federal Rules of Civil Procedure (FRCP) specify that “electronically stored information” (ESI) is a formal category of discoverable information

ESI is subject to production, and the management of ESI impacts eDiscovery costs and risks

There is a need to understand the “sources” of ESI

The word “preserving” appears in the rules for the first time

There will be an early “meet and confer”

Yet...

“Less than 20 percent of companies claim to be well prepared for eDiscovery issues” [1]

“47 percent of those surveyed say they’re only ‘somewhat confident’ that they could meet eDiscovery requests, while 10 percent are not at all confident” [2]

63 percent of respondents have been asked to perform a legal or compliance request—up from 57 percent one year ago” [3]

The average cost of a single eDiscovery case is \$1.5 million [4]

¹ eDiscovery Best Practices, Forrester Research, September 24, 2007

² Storage Magazine, November 2007

³ Storage Magazine, November 2007

⁴ Gartner Research, October 12, 2006



Effective information policy management for good corporate governance

- Properly discover, classify, and manage all information according to business value and risk
- Make appropriate decisions *before* litigation about the retention, deletion, and security of unstructured information
- Comply with corporate policy and regulations by knowing exactly what information is stored and where

Enables a quick and cost-effective response to discovery requests

- Crawl and classify active data on your network servers, storage systems, and personal computers without disrupting your end users
- Begin to analyze active data in hours or days versus weeks or months
- Improve your ability to cull down potentially responsive information

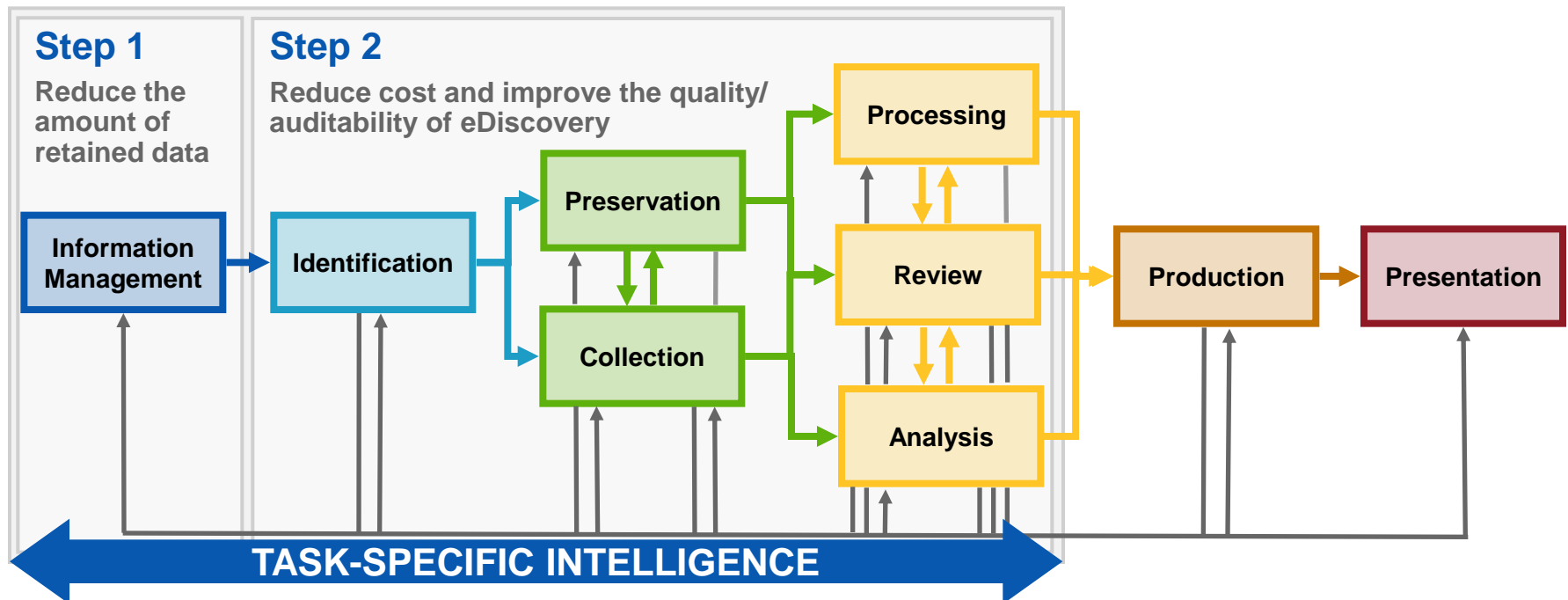


Electronic Discovery Reference Model (EDRM)

The Industry Model for Electronic Discovery

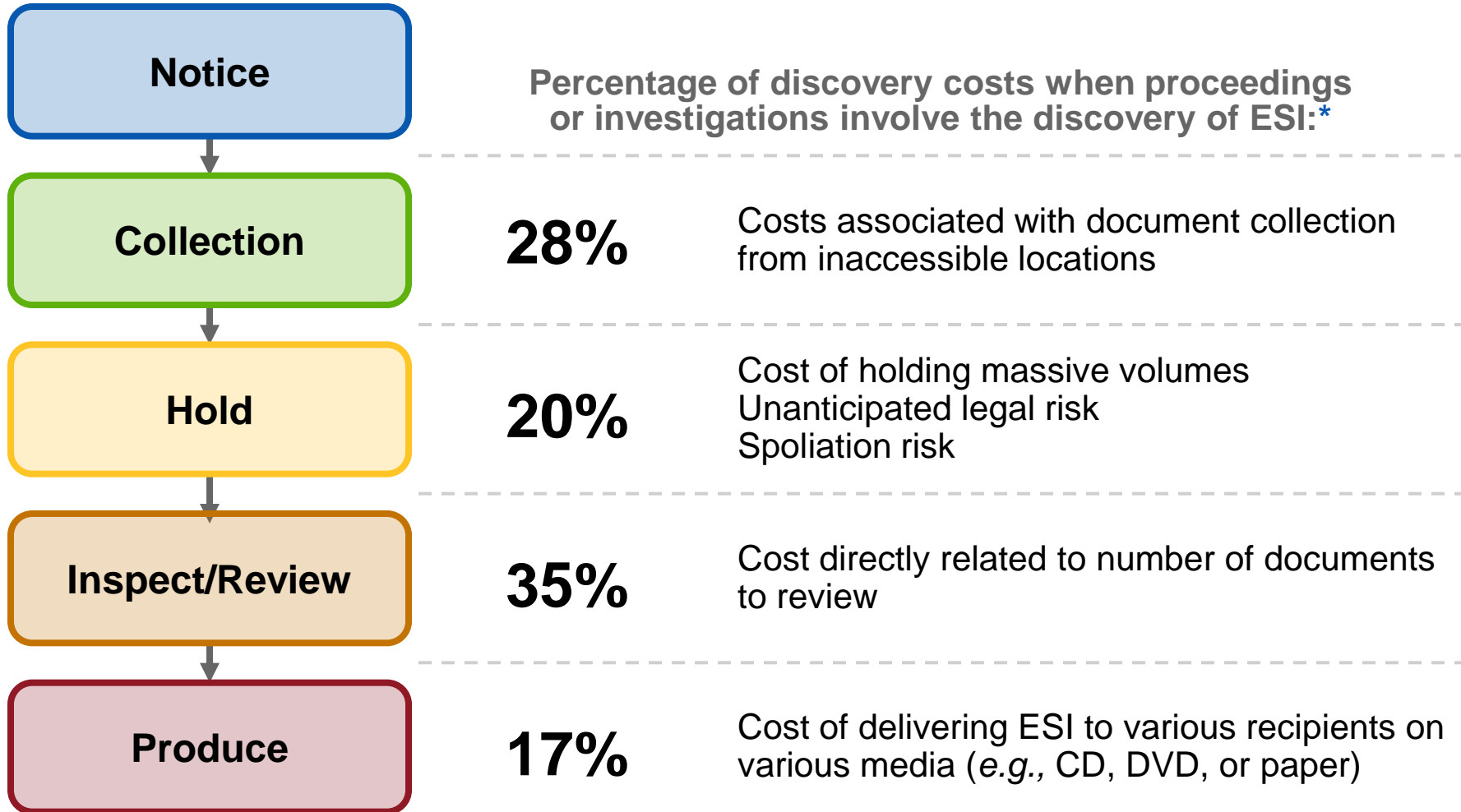
Technology can transform the traditional eDiscovery process

- Analysis can begin soon after solution is installed
- Two main steps to an effective eDiscovery process



An eDiscovery Request: A Repeatable Business Process

Common Subprocesses



* Source: Enterprise Strategy Group, 2007

eDiscovery Challenges

A Reactive, Time-Consuming, and Costly Endeavor

Business Challenges

Difficulty collecting, preserving, and producing unstructured ESI from disparate locations

Live, unmanaged, and unstructured data can be altered and deleted after “duty to preserve” exists

Massive amount of potentially responsive ESI sent for legal review

Collect and Hold

Cause

Many companies are experiencing vastly expanding ESI sources

Potentially responsive ESI not locked down and placed on “legal hold” in immutable storage

Limited in-house de-duplication and pre-review capabilities

Expanding ESI

Negative Impact

Identification, collection and preservation of ESI, is often reactive, time consuming, and costly

Failure to implement “legal hold” for litigation events can result in severe fines

High eDiscovery service provider fees associated with content review

Cost and Risk

What Customers are Looking For

Business Needs

Efficiency

- Reduce costs with more efficient in-house data collection and preservation

Speed to resolution

- Improve analysis and speed to resolution with faster preservation and processing

Downstream savings

- Create downstream savings by sending less content to outside counsel

Reduced risk

- Improve ability to implement and manage litigation holds on ESI in active systems

Technical Needs

Automation

- Reduce administration and overhead of disjointed eDiscovery tools

Integration

- Centralize collection from file, e-mail, and content management systems

Control

- Gain control of unmanaged and potentially responsive data in a secured infrastructure

Visibility

- Run sophisticated queries and out-of-the-box reports across unstructured data sources

EMC Solution for eDiscovery Collection

The EMC Solution for eDiscovery Collection automates the in-house collection and hold of unstructured content across live data sources

Transparent indexing and collection appliance

- Deployed very quickly
- Integrated into the EMC infrastructure

Enables corporate Legal and IT teams to:

- Proactively discover, classify, manage, and preserve information according to business value and risk
- Respond rapidly and efficiently to litigation, investigation, or compliance requirements
- Contain the costs of electronic discovery while mitigating organizational risk
- Employ lab-tested, defensible technology and eDiscovery best practices

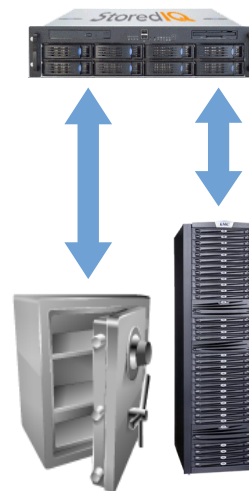
Solution Components

EMC Centera Governance Edition storage:
Optimized to manage the retention and disposition
of content stored by applications for internal
governance and regulatory requirements

- Immutable platform [write once, ready many (WORM) compatible]
- Content authenticity and integrity guaranteed
- Retention and disposition management capabilities
- Litigation hold

StoredIQ 4.5 appliance

EMC Consulting Services



Integrations

E-Mail Systems:

Microsoft Exchange

File Systems:

Desktops

Laptops

CIFS Fileshares

NFS Fileshares

Content Management Systems:

EMC Documentum

Microsoft SharePoint

Solution Capabilities

Provide for Greater Litigation Readiness: StoredIQ and EMC Centera

Begin to analyze active data in hours or days versus weeks or months

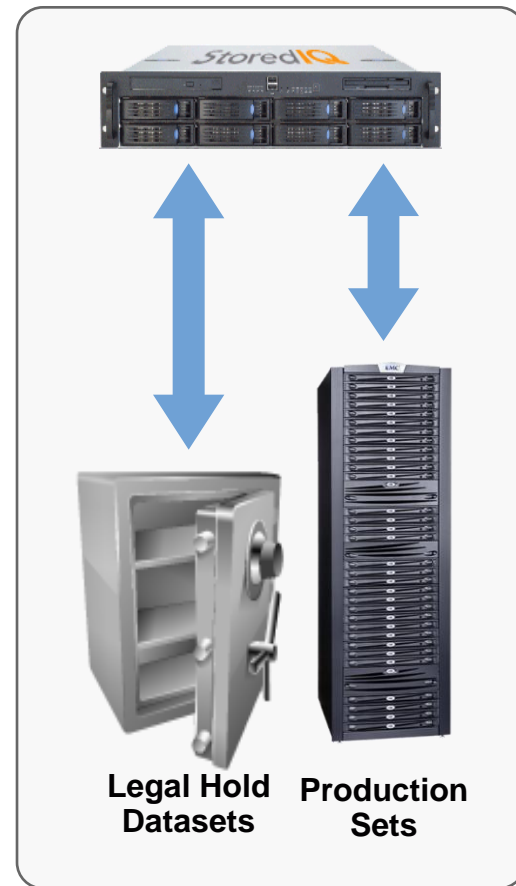
Automatic creation of load files supporting a variety of formats, including Lexis/Nexis Concordance

Act on individual files within containers such as PST, OST, NSF, ZIP, TAR, RAR, MSG, and EML files

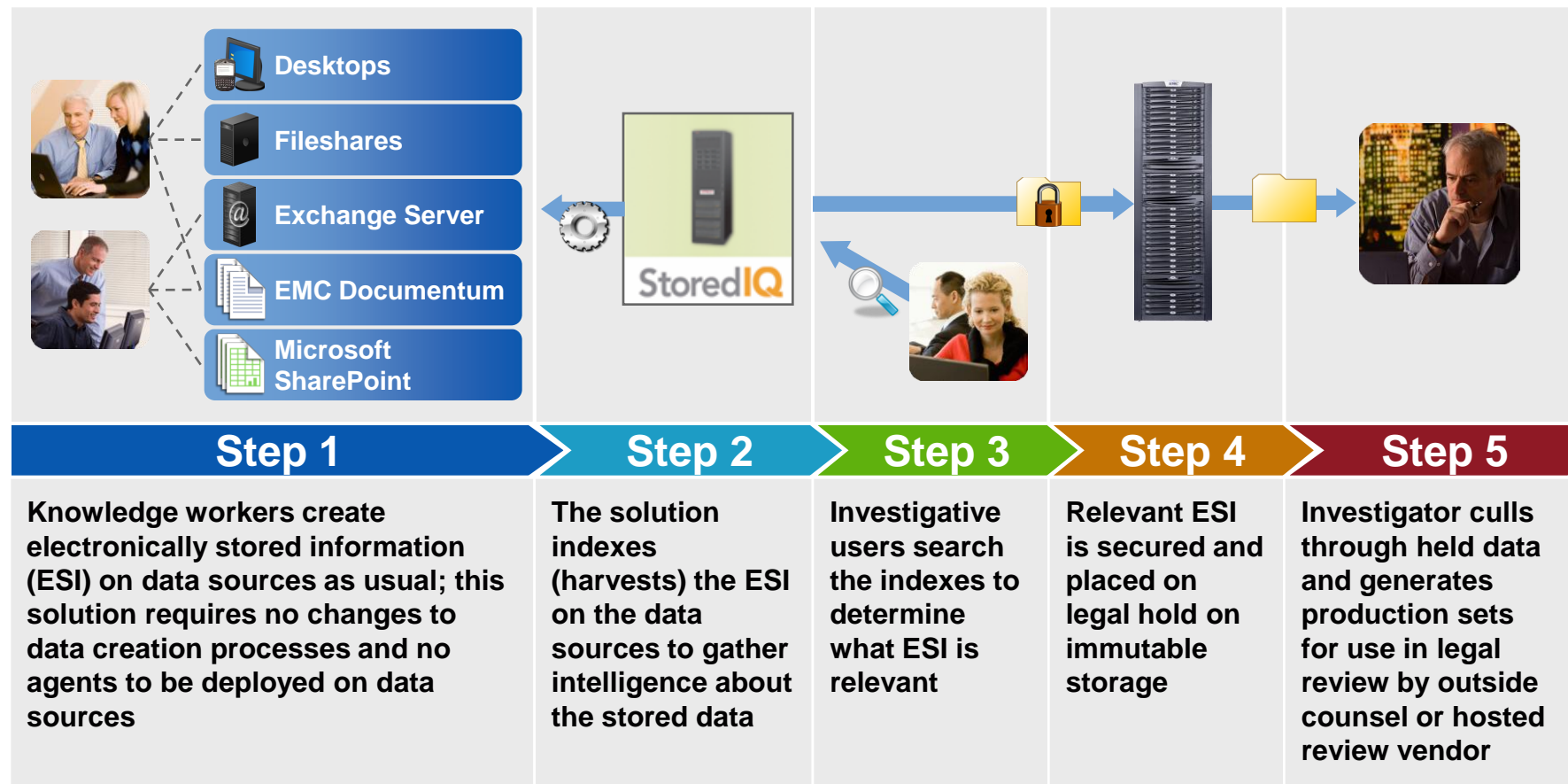
Advanced de-duplication capabilities

Recognize and process more than 300 different file types in their native format

Enable forensically sound and defensible eDiscovery processes

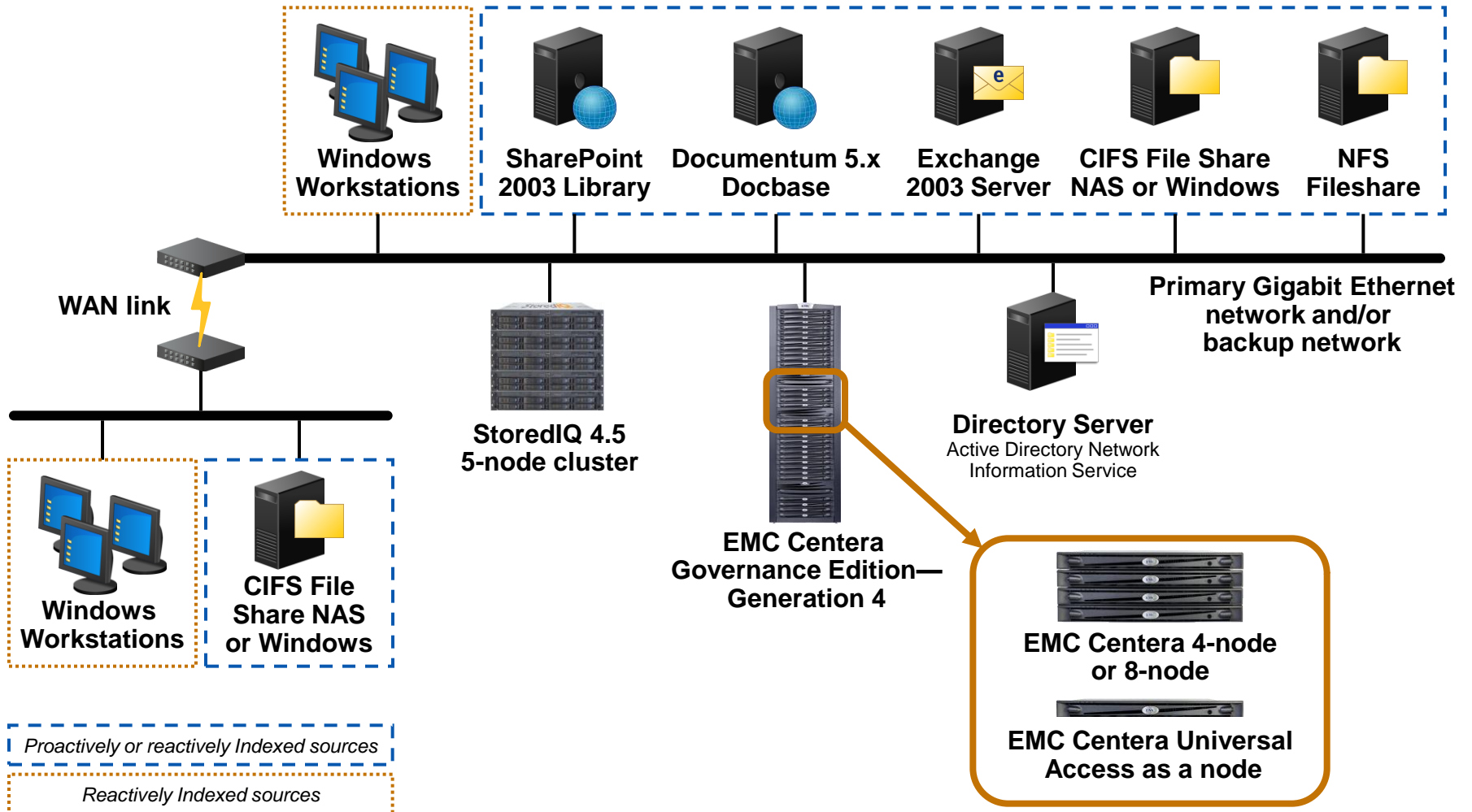


EMC Solution for eDiscovery Collection: Data Flow



Representative Validated Reference Architecture

Architecture to Support up to 20 TB of Managed Storage



EMC Solution for eDiscovery Collection: Process Flow

- ① Awareness
- ② Early Case Assessment
- ③ Preservation
- ④ Data Topology Reporting
- ⑤ Meet and Confer
- ⑥ Review Preparation

EMC Solution for eDiscovery Collection: Process Flow

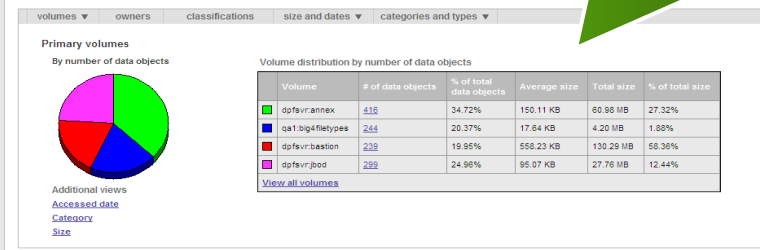
2 Early Case Assessment

Identify relevant Custodians and associated data sources

- Review data corresponding to Custodians from identified data sources
- Identify the scope and amount of data to be collected and preserved

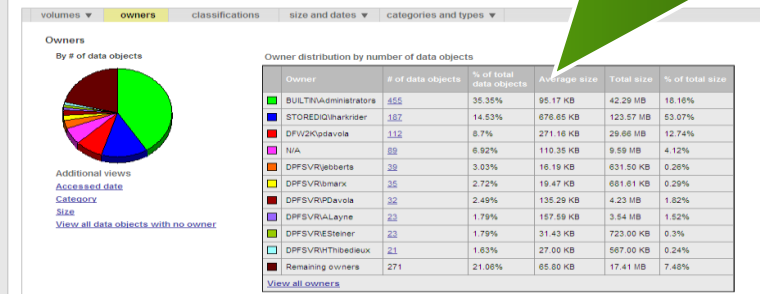
Data Source View

data explorers



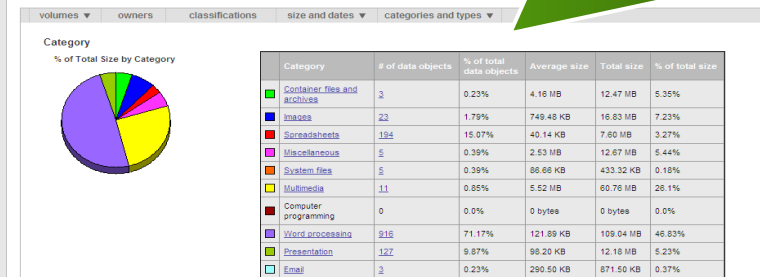
Custodian View

data explorers



Category and Type Views

data explorers



2 Early Case Assessment (continued)

Query-based analysis

- Based on attributes and/or full text
- Identify all relevant Custodians and Custodian-relevant data in shared data locations

1. Attribute Categories and Samples

The screenshot displays the 'Browse for attributes' section of the EMC eDiscovery interface. On the left, a tree view lists various attribute categories, with 'System metadata' highlighted. On the right, a list of attributes is shown, including 'Accessed date', 'Container path', 'Created date', 'Extension', 'Group', 'Modified date', 'Object name', and 'Owner'. A green callout box labeled '1. Attribute Categories and Samples' points to this list.

Below the attribute list, a 'Configure a condition' dialog box is open. It shows a 'User-defined attribute' set to 'GB - Contract Declarations'. The 'Operator' is set to 'appears' and the 'Value' is set to 'at least 2 times'. A green callout box labeled '2. Attribute Evaluation' points to this dialog. To the right of the dialog, a text box provides instructions: 'When choosing number of occurrences: • When entering numbers, enter in a number greater than or equal to zero with no decimal points. • In the 'between' case, make sure the first number is less than the second.'

At the bottom of the interface, the 'expression editor' is visible. It contains a visual representation of a Boolean-based attribute expression: 'Contract Lang Or Contract Struct And Signatory And Genl Contrac'. A green callout box labeled '3. Boolean-Based Attribute Expressions' points to this editor. The editor includes buttons for 'Return to editor', 'Add parentheses', and 'drag & delete parentheses'.

2 Early Case Assessment (continued)

Use e-mail attributes to
reconstruct threads

Use the deep attribute
and content query
functionality to identify
relevant Custodians and
Custodian-relevant data
in shared data locations

The screenshot displays the 'current expression' configuration window. At the top, it shows 'Set 1' and 'Set 3' connected by an 'And' operator. Below this, the 'Set 3' configuration is shown with the text 'Find only data objects that match: Any of the following'. A table lists the conditions:

Category	Attribute	Operator	Value	Applies to
Email	To	is	Tom Smith	Single data objects
Email	To	is	Doug Fierro	Single data objects

Below the table, there is a 'Define a condition:' field and a 'Browse' button. A 'Browse for attributes' dialog is open, showing a tree view of categories. The 'Email' category is selected, and a list of attributes is shown on the right, with 'To' highlighted.

EMC Solution for eDiscovery Collection: Process Flow

2 Early Case Assessment (continued)

Collection/preservation tagging

Verify your search results through review of the relevant metadata and e-mail/document content

The screenshot displays the EMC eDiscovery interface with three main sections:

- Result Set:** A table titled "attribute search results" with columns for Object name, System path, Accessed date, Modified date, and Owner. The table lists several PDF files, with "s_Option_Agreement.doc" highlighted in orange.
- Metadata View:** A detailed view of the selected document's metadata, including System metadata (Type: Microsoft Word, Size: 40.50 KB, Owner: STOREDQI@har), Attributes (Classifications: matter03109, MINDEX, processing 03109), Query tags (shares w/2 50000, (stock w/2 option) AND g...), Entities (GB - Form Blank, Dates, GB - EE Rating Terms, First names (2ND), GB - Contract TOC, Companies, GB - EE Ora Terms), and Data object properties (App Name, Author, ByteCount, ObjCount, StatusCode, Create_DTM, EditTime).
- Pure HTML Content Viewer:** A view of the document's content, showing a "Form 6-5 INCENTIVE STOCK OPTION AGREEMENT" with explanatory notes and terms. The text includes details about the agreement between Aquarius a Nevada corporation and Sykes, including the option price and grant date.

Blue arrows indicate the flow from the search results to the metadata view, and then to the content viewer. Green callout boxes label each section: "Result Set", "Metadata View", and "Pure HTML Content Viewer".

EMC Solution for eDiscovery Collection: Process Flow

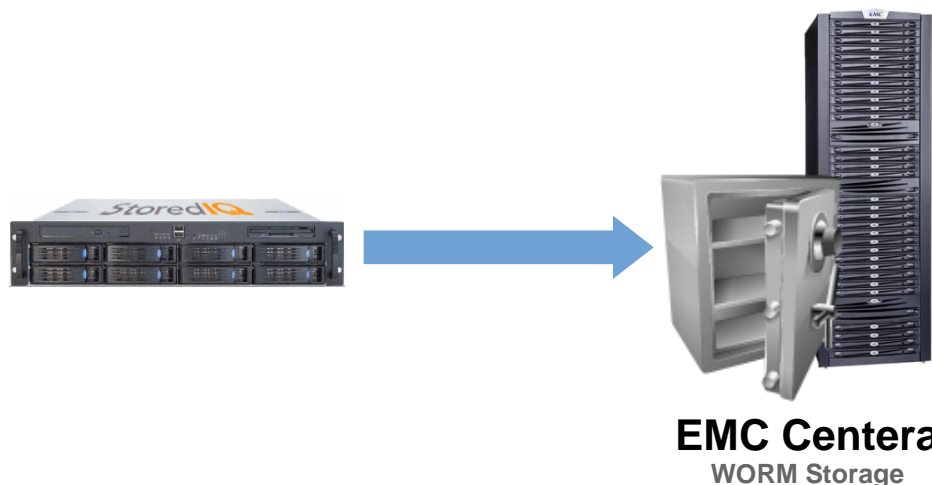
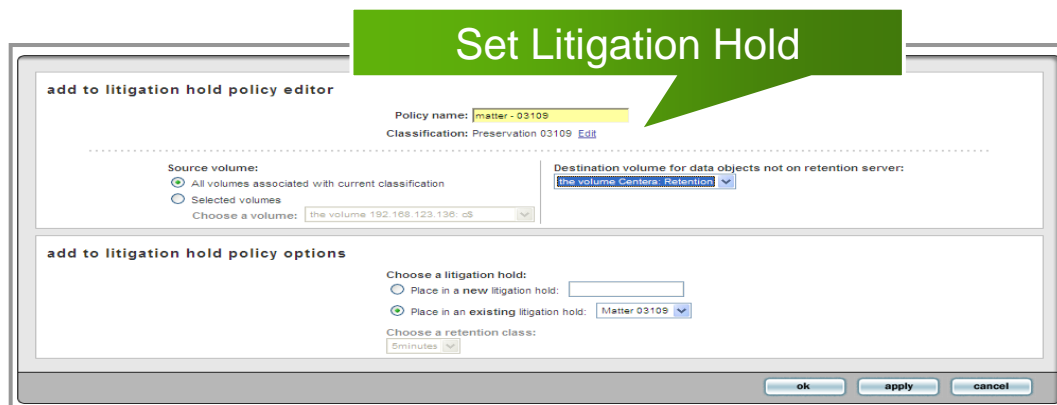
③ Preservation

Simple “one-click” process

Chain of custody and
authenticity preserved

No alteration of metadata

Original, full-object path
information is properly
maintained

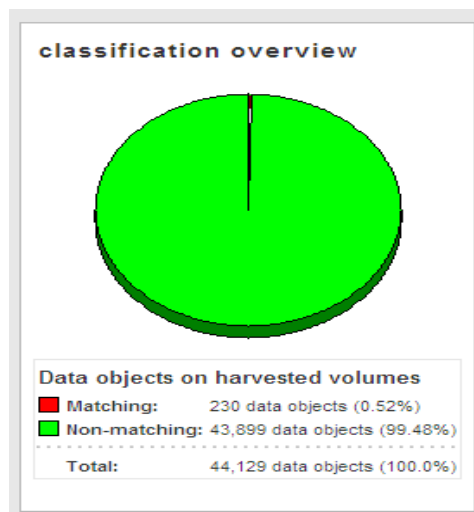


4 Data Topology Reporting

Create data topology for potentially responsive data with full analytic reporting

5 Meet and Confer

In preparation for the “meet and confer,” generate data topology maps and query reports for all data under preservation



**Preservation Data
Topology Map/Reports**

StoredIQ
Data topology map

Classification: Preservation 03109

Generated: 2007-12-02 05:29 PM

Locations surveyed for this report

Volume	Owner	Category	Modified date	Accessed date	# of data objects	# of duplicate objects	# of unique objects	Size of unique objects
192.	BUILT	Container file	This week	This week	1	0	1	290.50 KB
192.	DFW2	Email	1 to 4 weeks ago	1 to 4 weeks ago	0	0	0	0 bytes
192.	DFW2	Images	1 to 4 weeks ago	1 to 3 months ago	222	113	109	18.11 MB

EMC Solution for eDiscovery Collection

Process Flow

⑥ Review Preparation

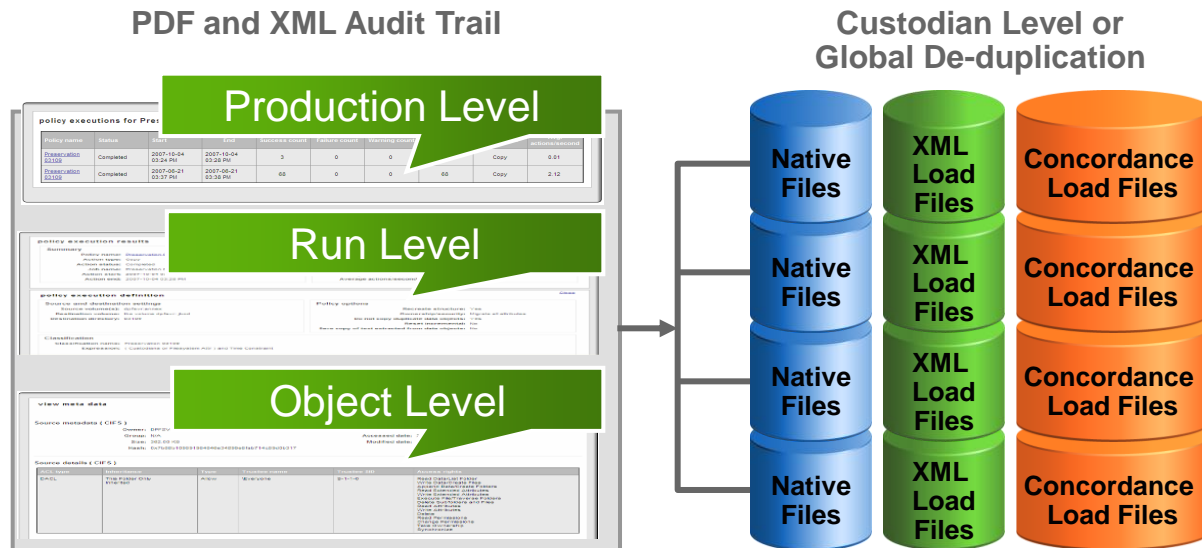
Generate a review load file of all responsive items negotiated in the meet and confer

De-duplication

Container file capabilities

Automatic prefixing

Automated rolling productions



EMC Select Partner Program

StoredIQ: An EMC Select Partner

Program extends EMC's end-to-end eDiscovery offering via complementary third-party portfolio

EMC Select products are qualified for interoperability with EMC technology

- EMC E-Lab tested where applicable

EMC Select vendors deliver service and support

- Vendors sign joint Customer Support Agreements with EMC for joint escalation and efficient resolution of problems

EMC's one source for end-to-end information infrastructure solutions



& **StoredIQ**



Expense versus Investment: Complementary EMC Products

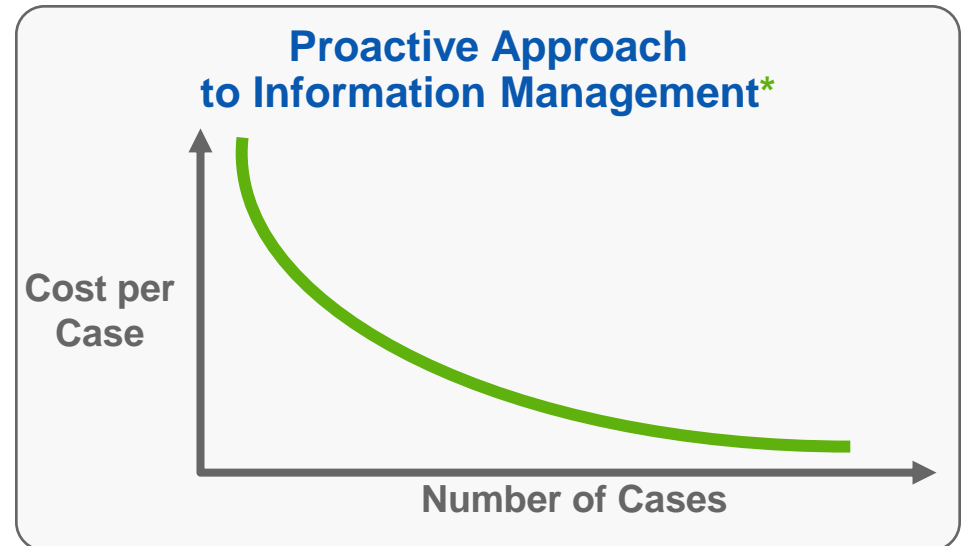
For Proactive Information Management

EMC Documentum

- Records Management
- Retention Policy Services
- Information Rights Management
- Secure Collaboration

EMC EmailXtender

- Stores electronic communications in a centralized archive
- Allows you to quickly find and retrieve relevant messages



eDiscovery is often cited as the primary motivator behind investments in the records and information infrastructure category

*Source: IDC "Worldwide Compliance Infrastructure 2007-2011 Forecast: Compliant Information Infrastructure, Data Privacy, and IT Risk and Compliance Management Underpin Spending"

EMC²[®]

where information lives[®]