



Backup & Recovery Considerations for the EMC Documentum Platform

By Mike Fernandes, Vice President, Products, enChoice

The EMC Documentum enterprise content management platform is a powerful system enabling organizations to build robust business applications. Companies invest significant resources into Documentum to streamline critical processes, optimize operations and facilitate compliance with corporate and industry regulations. These applications are intricately linked to massive, ever-growing amounts of business-critical information.

Protecting this information from corruption and loss is vital, however, effective backup and recovery can be challenging due to stringent service levels and the complex relationships between content and metadata. This white paper discusses these challenges and how they can be addressed with application-aware business continuity solutions.



Table of Contents

Overview	3
Challenge 1: Meeting business requirements for Documentum availability during system backups	3
Why “hot” Documentum backups can cause data loss and corruption	3
CYA HOTBackup provides online Documentum platform backup	4
Challenge 2: Recovering content and metadata without a full system restore	5
Traditional enterprise backup systems do not provide granular recovery	5
Why it’s difficult to recover from operational (logical) loss	5
The importance of metadata	6
CYA SmartRecovery provides hot, granular recovery	6
Why an application-aware solution is important for granular recovery	6
Challenge 3: Reducing data loss from hours to minutes	7
CYA: The ultimate recovery solution for EMC Documentum	8
Summary	8
About enChoice	8

Overview

Organizations invest significant resources into the EMC® Documentum® enterprise content management (ECM) system to realize business efficiencies and improve fiscal performance. Applications such as Documentum Compliance Manager, Digital Asset Manager, Records Manager and Business Process Manager provide the framework and capabilities companies need to streamline processes, optimize operations and facilitate compliance with corporate and industry regulations.

Regardless of the applications deployed, the information stored within the Documentum repository is a complex network of metadata and content. Because these applications drive core business initiatives, it is critical that every organization have a strategy in place that ensures full recovery from all causes of information loss.

System failures, site disasters and common operational incidents such as human and programmatic errors can cause the loss or corruption of one, several or thousands of pieces of information. This often results in significant, unacceptable consequences including lost revenue, a damaged reputation, and in extreme cases, a complete business shutdown.

A 2013 study conducted by the Ponemon Institute found that the average cost of downtime across industries was approximately \$7,900 per minute, with an average reported incident length of 86 minutes—a cost of \$690,000.¹ And while the aggregate costs of data loss incidents are hard to quantify, Vanson Bourne's 2013 Global IT Trust Curve survey found that they cost businesses close to \$586,000 in revenue loss per year.² Additional consequences of these incidents were diverse and highly damaging, and included loss of employee productivity, loss of business to competitors, and loss of customer confidence/loyalty.

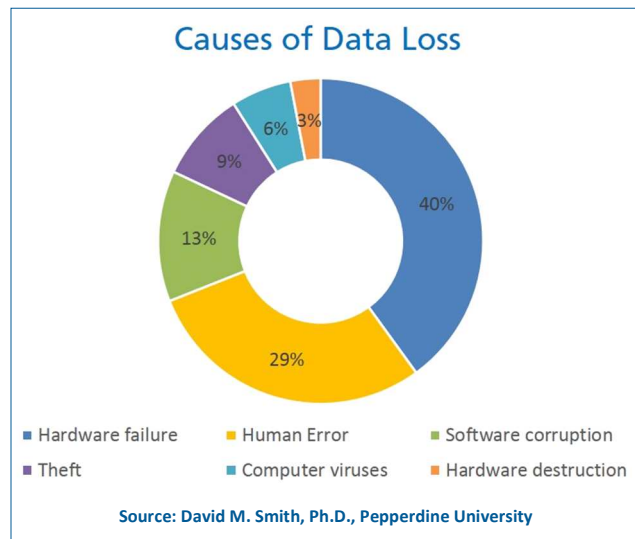
Unfortunately, traditional enterprise backup and recovery strategies do not fully protect Documentum system information – in fact, they often leave information unrecoverable or in an inconsistent state upon recovery. This document outlines the specific challenges associated with recovering Documentum repository information and provides an overview of the solutions that enable an effective and complete recovery strategy.

Challenge 1: Meeting Business Requirements for Documentum Availability During System Backup

In today's era of global, "always on" operations, businesses require continuous access to information for worldwide employees, customers and partners. This requires IT organizations to provide a high level of service for application and information availability.

Why "hot" Documentum backups can cause data loss and corruption

With most applications, technologies such as snapshots or "hot" backup solutions make it relatively simple to provide high availability. However, these technologies are not suitable for Documentum as they back up underlying Documentum repository components at different times and therefore cannot ensure consistency between the components. For example, the database is usually much smaller than the content storage location,

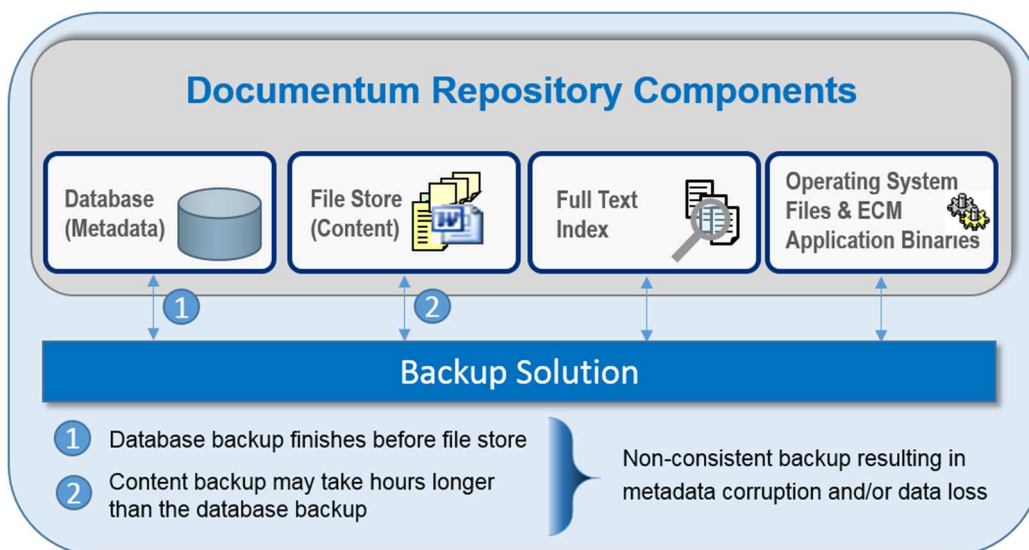


¹2013 Cost of Data Center Outages, Ponemon Institute

²2013 Vanson Bourne IT Trust Curve Survey

so the database backup is usually completed much sooner than the content backup. As a result, metadata may exist for content that's not present on the file system, or content may exist for which there is no metadata.

Deploying traditional backup technologies without care can result in corruption or information loss if they need to be used to recover the Documentum repository. For this reason, "hot" backups are not recommended or supported by EMC Documentum as described in EMC support note esg1224.



Synchronization issues with hot backup solutions

Traditional backup solutions offer database and open file agents or modules that enable a hot backup for each component, but they cannot ensure consistency between the database and content.

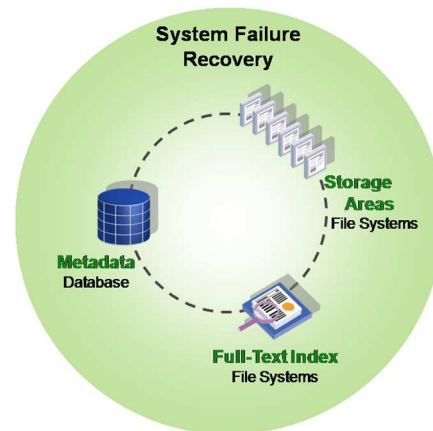
CYA HOTBackup provides online Documentum platform backup

Leading recovery solutions provide modules that enable a database and file system backup to be performed while the application is online. However, as previously described there is no coordination or consistency between these backups, which leads to data corruption.

CYA HOTBackup™ from enChoice addresses these consistency issues. The software solution enables companies to mitigate and even eliminate planned Documentum downtime by performing hot, consistent backups of the entire Documentum application that are automatically stored within EMC NetWorker, EMC Avamar, Symantec NetBackup™, IBM® Tivoli Storage Manager (TSM) or HP Data Protector.

CYA HOTBackup targets all Documentum application components encompassing the content storage area, index files and database files. This simplifies the Documentum backup strategy by allowing synchronized backups to be performed "hot" while the application remains online.

It also ensures that the Documentum repository can be recovered in its entirety after system failures and disasters. CYA HOTBackup can perform a Documentum repository backup as often as every 24 hours, resulting in a 24 hour data loss window.



Common components backed up with CYA HOTBackup

Challenge #2 – Recovering Content & Metadata Without a Full System Restore

The days when a daily backup provided adequate protection for critical business processes are long gone. Today, downtime - *any* downtime - translates to real and significant costs: in lost productivity because employees can't work, lost business because orders can't be taken, and customer dissatisfaction as your downtime translates to their downtime.

A study by Infonetics Research found that overall downtime costs averaged an astounding 3.6% of annual revenue.³ Well over half of that downtime is unplanned. Of that unplanned downtime, 80 percent is due to operational incidents such as software and user errors, unknown and irreproducible bugs, accidental deletions or configuration changes, batch jobs that corrupt important data, and administrative errors.⁴

Traditional enterprise backup systems do not provide granular recovery

Traditional backup software and solutions such as RAID configurations, replication and snapshots are designed to protect systems from a complete loss. They are not well-suited for addressing situations where one, several or thousands of documents have been lost or corrupted in response to operational incidents.

- RAID is completely useless for operational recovery – if the metadata becomes corrupted, RAID merely ensures that the corrupt data is safe from disk failure.
- Tape backups and even snapshots are all-or-nothing solutions – they cannot prevent the loss of hours, or even days worth of data prior to the operational incident.
- Online replication systems require resynchronization from a backup server and will replicate corruptions to the secondary server.

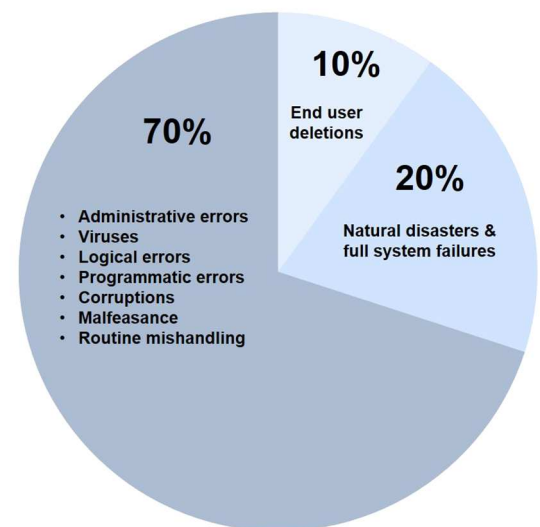
All of these technologies have their place in an effective overall application protection strategy. However, none of them address recovering content and metadata at a granular level in an EMC Documentum repository without an offline system restore to a prior point in time. If a company has to roll back its entire Documentum system to its last known good state to recover information, it will lose all the work employees generated after that point in time. In addition, because Documentum must be taken offline for the rollback, significant productivity is lost and transactions cannot be processed.

As a case in point, in the life sciences industry, it's anticipated that for every day an ECM application is offline, there's a daily cost to the organization of \$1 million due to lost productivity, halted test process cycles and delayed market availability of products.⁵

The primary business driver behind the implementation of Documentum is often to gain efficiencies within business processes and ultimately improve fiscal performance. System failures, user mishandling and any other occurrence that result in lost metadata puts that efficiency goal in jeopardy.

Why it's difficult to recover from operational (logical) loss

Documentum dynamically creates and manages the entire content lifecycle along with its associated metadata. Metadata is comprised of audit trails, digital signatures, version information, renditions, security, user information and custom attributes among other properties. In Documentum, each one of these properties is related to an



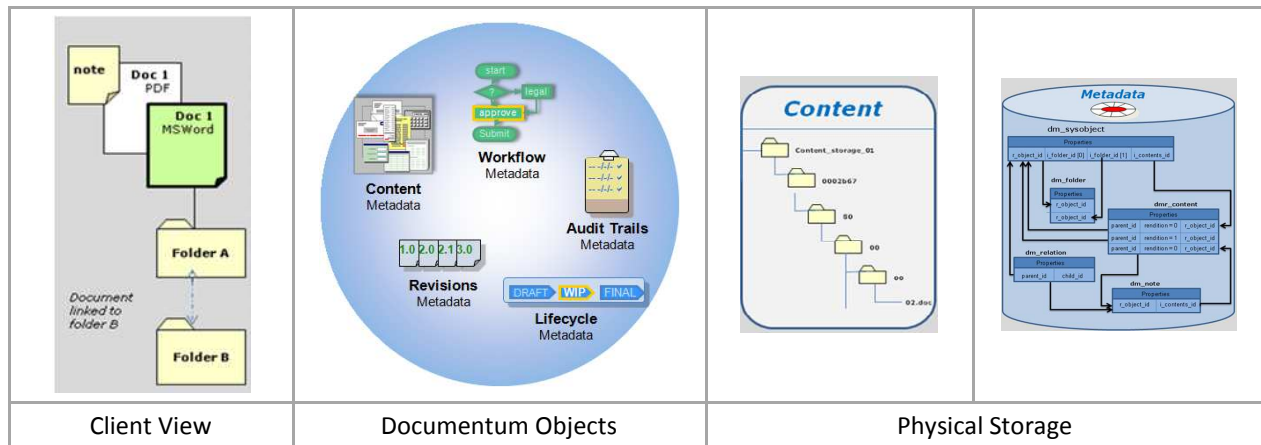
Causes of ECM data loss
Source: AIIM International

³ The Costs of Enterprise Downtime, North America, Infonetics Research

⁴ AIIM International & Strategic Research Inc.

⁵ Tufts Center for the Study of Drug Development

object in the Documentum Content Server Object Model. Each property is also associated with content stored on the physical storage device using a hexadecimal naming convention. This creates extreme complexities and difficulties in recovering documents and their related properties.



Documentum has a complex metadata structure

The importance of metadata

Metadata is defined as “data about data.” One of the core values of an ECM system is its ability to maintain extensive amounts of complex metadata related to the various documents it manages - data that can be just as crucial as the documents themselves. This is especially true as companies strive to comply with the various regulations that are specific to the preservation and recoverability of authentic metadata, including but not exclusive to the Sarbanes-Oxley (SOX) Act of 2002 for all public companies, Food and Drug Administration (FDA) regulations, Security and Exchange Commission (SEC) policies for the financial services industry, and the Federal Executive Branch Continuity of Operations (COOP).

Furthermore, as it relates to eDiscovery, the ability to accurately recover metadata is just as important—if not more so—than the ability to recover business applications and documents. For example, consider the Discovery Rules Amendment of the Federal Rules of Civil Procedure, effective December 1, 2006, which stipulates that companies involved with litigation must, by default, not only produce all requested documents and e-mail in their original electronic state, but must also produce the accompanying unaltered electronic audit trail. Companies also must be able to demonstrate that the audit trail was not tampered with.

CYA SmartRecovery provides hot, granular recovery of content & metadata from operational incidents

CYA SmartRecovery™ from enChoice is the only solution enabling rapid recovery from all types of operational incidents. Just one administrator can restore one, several, or thousands of objects (content and associated metadata) back to the repository in their original states while Documentum remains online. There is no need to recreate lost/corrupt information, or incur additional data loss by performing a full system restore to recover granular objects. In addition, CYA SmartRecovery ensures compliance with regulations requiring the preservation of information (i.e. 21 CFR Part 11) by recovering metadata with original date and time stamps.

Why an application-aware solution is Important for granular recovery

In order to ensure complete recovery from operational incidents, it is critical to use an application-aware solution. This means that the solution is aware of the Documentum object model, and the relationships that objects have (or should have) to each other within that model.

Application-aware solutions are uniquely able to:

- **Capture and restore content AND the metadata that makes it useful (i.e. permissions, attributes, annotations).** The solution must “understand” and maintain the relationships between content and metadata to ensure a successful backup and recovery. CYA SmartRecovery recovers the content and metadata components in their original state and context.
- **Check for key object relationships and identify broken relationships during the capture process.** This reduces the potential of capturing metadata corruptions.
- **Proactively scan the repository.** CYA SmartRecovery performs over 30 checks for key object relationships, identifies broken relationships, and categorizes corruptions by severity. This information enables the administrator to prioritize and build a resolution plan for addressing the corruptions.
- **Adhere to records retention policies defined within the Documentum repository.** An application-aware solution such as CYA SmartRecovery ensures that information is destroyed appropriately from CYA backups.

CYA SmartRecovery’s primary benefits for Documentum deployments are:

- Enables organizations to quickly restore one, several or thousands of objects back into the repository to their original states
- Safeguards complex lifecycle information against loss due to programmatic and logical errors, malfeasance, human errors, deletions and viruses
- Facilitates accurate, rapid response to regulatory audits, inquiries, inspections and/or eDiscovery

Challenge #3 – Reducing Data Loss from Hours to Minutes

When organizations need to reduce data loss from weeks or hours to minutes, the only solution that can meet the challenge is CYA SmartRecovery.

CYA SmartRecovery’s application-aware technology captures metadata and content through the Documentum Content Server APIs in a single pass while the application is online. Because it only captures incremental changes and has transparent repository impact, it can be scheduled to run as often as every fifteen minutes, enabling companies to achieve stringent recovery point objectives (RPOs).

In addition to ensuring consistency between content and metadata, this approach verifies the integrity of the information to ensure a “clean” capture with no corrupt data.

Industry regulations and guidelines requiring data preservation

Life Sciences

- FDA CFR Part 11
- FDA CFR Parts 210/211 (cGMP)
- ICH Guideline E6
- EU Commission Directive 2003/94 (Good Manufacturing Process Guidelines)

Financial Services/Insurance

- The Dodd-Frank Wall Street Reform and Consumer Protection Act
- Sarbanes-Oxley
- Securities and Exchange Commission (SEC) Rule 17a-4
- FINRA “Customer Suitability” rules
- NAIC audit rules

Energy

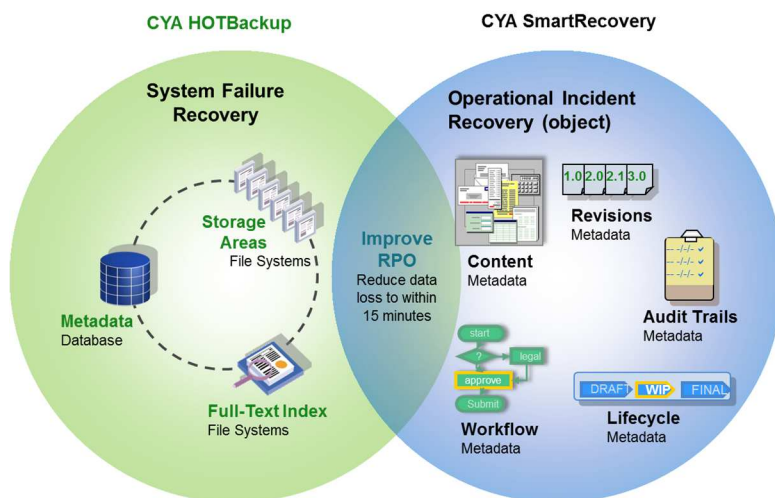
- North American Electric Reliability Council (NERC) 1200 (1216.1)
- Federal Energy Regulatory Commission (FERC) RM01-12-00 (Appendix G)
- NERC Security Guidelines for the Electricity Sector
- Occupational Safety and Health Administration (OSHA) Regulation 1910.119 for Process Safety Management (PSM)
- ISO 9000

.....and many more

CYA: The Ultimate Recovery Solution for EMC Documentum

CYA SmartRecovery and CYA HOTBackup combined with traditional enterprise backup and recovery systems provides the “ultimate recovery solution” for EMC Documentum. System point-in-time backups with enterprise systems including EMC NetWorker, EMC Avamar, Symantec NetBackup, IBM TSM and HP Data Protector work seamlessly with CYA software to provide complete protection from system failures, disasters and operational incidents.

The Documentum repository database and content files are restored from the last point-in-time backup performed by the enterprise system, and subsequent CYA incremental backups are applied for a complete restoration of the repository to the point of the last CYA capture, reducing data loss to as little as 15 minutes.



CYA HOTBackup and CYA SmartRecovery seamlessly leverage enterprise backup solutions to recover the entire Documentum system and reduce data loss to 15 minutes.

Summary

After investing significant resources into an EMC Documentum system to achieve business efficiencies and a bigger bottom line, it's critical to protect that investment by ensuring 24x7 availability and end-to-end protection against all forms of data loss and corruption. This can be a daunting challenge due to the complexity of the Documentum platform, which consists of a complex environment posing special backup and recovery challenges that cannot be addressed with traditional enterprise backup systems.

CYA SmartRecovery and CYA HOTBackup from enChoice uniquely enables companies to fully protect their Documentum investments by working with traditional backup systems to enable hot, application-aware backup and recovery. CYA software enables companies to overcome Documentum backup and recovery challenges by:

- **Enabling hot system backups** in tandem with traditional disaster recovery solutions
- **Providing hot, granular backup and recovery** for Documentum objects (content and metadata)
- **Reducing information loss** to just 15 minutes to achieve stringent service levels
- **Facilitating compliance** with regulatory mandates and legal requirements calling for data preservation

To learn more about CYA software visit enchoice.com, email cya@enchoice.com, or call +1.480.477.3838.

About enChoice

enChoice®, Inc. provides enterprise content management software solutions that improve time-to-solution, the user experience and backup and recovery service levels. enChoice software solutions are complemented by end-to-end ECM services and a unique “single point of contact” support program. The company is recognized industry-wide for its innovative, application-aware solutions for EMC Documentum and IBM FileNet. Founded in 1993, enChoice is an EMC Technology Connect Select Partner.

