Contents

Challenges of Content Archiving ..........................................................................................1
Open Text Archiving for Content Server Architecture .......................................................2
  Content Server Storage Provider ......................................................................................2
  Archive Server Storage Provider .....................................................................................5
  Content Move ...................................................................................................................5
Intelligent Storage Management..........................................................................................7
  Content Value versus Content Cost ...............................................................................7
Challenges of Content Archiving

Traditional document management systems are an excellent and proven solution for supporting collaborative project work, eliminating shared drives, and facilitating document review and approval. They enable easy corporate content retrieval and sharing across departments, divisions, and even organizations.

However, when managing huge quantities of documents that must be retained according to very strict retention plans, traditional document management systems face certain and inherent limitations. Simply connecting a document management system to storage hardware isn’t always sufficient. Compliance with regulations often means that organizations need a solution for managing content throughout the entire Enterprise Content Management (ECM) lifecycle, from creation through publication, to long-term storage and eventual destruction.

Open Text Archiving for Content Server provides archiving capabilities for systems based on Open Text Content Server. Open Text Archiving for Content Server is included in Open Text Content Lifecycle Management and can be added as an option to other systems based on Content Server, such as Open Text Document Management.

Open Text Archiving for Content Server provides:

- Access
  - Long-term content availability
  - Long-term content readability
- Security Services
  - Tamper-proof storage
  - Access control
  - Encryption
  - Controlled deletion
  - Auditing
- Cost Savings
  - Single instance archiving
  - Data compression
- Administrative Benefits
  - Backup and restore
  - Disaster recovery
Open Text Archiving for Content Server Architecture

Open Text Archiving for Content Server enables Content Server-based systems to use Open Text Archive Server for a variety of storage options beyond the External File Store (EFS) and internal database binary large object (BLOB) storage provided by Content Server alone.

Archiving for Content Server comprises Open Text Archive Server and Open Text Archive Server Storage Provider. The following diagram illustrates how these components connect to Content Server and Archive Server. Content Server content (from both the file system and the database) can be moved to Archive Server. Archive Server virtually stores the content, maintaining complete independence from the physical storage system.

Content Server Storage Provider

Open Text Content Server, shown in the diagram above, includes its own “Content Server Storage Provider” separate from the “Archive Server Storage Provider.” When Archiving for Content Server is deployed, administrators can use the Content Server Storage Provider to configure storage rules. These rules control which Content...
Server objects, (such as documents and folders) are archived in Archive Server, and which are stored in specific logical archives.

In Figure 2 below, the expression column describes the rules established for selectively moving content to the storage medium. In this specific example, content labeled with the “SOX” category is stored within the Logical Storage Provider called “Sarbanes Oxley”, and content labeled with the “FDA” category is stored within the Logical Storage Provider called “FDADocuments.”

The separation between Content Server and Archive Server, combined with the virtualization of storage provided by Archive Server, means that Content Server does not require the technical information corresponding to "Sarbanes Oxley" and "FDADocuments" physical storage media. Separation allows for adjustment and upgrades to physical storage media, either because of changed regulations or technological advances, without requiring updates to archival rules.

Figure 2:
Storage rules specify what content is archived and to where

The Content Server Storage Provider component enables various storage providers to be defined in Content Server. Rules can then be established to control what content is written to which provider.

Figure 3:
Logical archives enable administrators to manage archived content selectively
The actual storage location (the Logical Storage Provider) of a document is fully transparent to end-users.

Figure 4: The versions page of a document in Content Server indicates the storage provider.

Figure 5 illustrates the functionality of Content Server Storage Provider: configure storage, establish and invoke storage rules, and move content to Archive Server.

Figure 5: Content Server Storage Provider enables storage rules to govern when content is written to Archive Server.
Archive Server Storage Provider

Archive Server Storage Provider is the component of Archiving for Content Server that establishes a connection between Content Server and Archive Server. The Archive Server Storage Provider allows the Content Server Storage Provider to use a logical archive as a storage provider.

All logical archives configured in Archive Server are shown in a pre-selection box of the Archive Server Storage Provider configuration page, and any of them can be used as a storage provider. Based on the “known-server” concept of Archive Server, even logical archives in an Archive Network are available for selection.\(^1\)

Archive Server Storage Provider supports signed URLs to prevent unauthorized access to Content Server documents on Archive Server.

Content Move

While the Storage Provider Rules indicate what content is to be stored and where, the Content Move module indicates when existing content should be moved from one Storage Provider to another. These are called “Content Move Jobs.”

With a Content Move Job, administrators can easily move any amount of content in Content Server from one Storage Provider to another. For instance, content within an External File Store (EFS) can be moved to a Logical Archive provided by Archive Server. With Content Move, administrators can migrate all of the content in an EFS to Archive Server, and then stop using the EFS altogether. Alternatively, administrators can apply move rules to a given set of documents within the EFS. Applying rules to a subset of EFS documents facilitates short-term EFS storage, while enabling use of Archive Server as a long-term storage repository. Content Move is not limited to moving content from an EFS; the tool can move documents and other objects from any Storage Provider.\(^2\)

\(^1\) For more information about Archive Networks please see the “Open Text Archive Server (Why Archiving Matters) Whitepaper”

\(^2\) There may be restrictions on the content and/or physical storage device used that are not compatible with Content Move.
Intelligent Storage Management for Open Text Content Server

Figure 6: Using Content Move, users first define rules, and then define the set to be moved and the time to move it.

Throughout the whole process, Content Move treats every version of a document as a single item, which means that different versions of documents and other content can be stored on different storage providers.

Content Move has an extended rule set. Users can define rules for how documents are moved, specify which documents to move, and schedule jobs to occur daily, weekly, monthly, and so on. Move operations are performed automatically, according to the specific rules set by the user. The move operation is tracked and logged in the content’s audit trail. The administrator is notified of any failures that occur as a result of hardware faults.
Intelligent Storage Management

Over time, archived records may be re-classified, reflecting both changes in content and/or regulatory requirements. These changes may require that content be moved to a different storage provider, and possibly a different storage medium. An archival system must be able to meet these changes without affecting the applications and users that might require content access. With the virtualization of storage provided by Archive Server and its robust and comprehensive policy engine, Open Text Archiving for Content Server ensures appropriate use of storage resources over the lifecycle of content.

Content Move provides special treatment for items under the control of Open Text Records Management. Records Management changes the properties of an item and automatically places an ID for each item in a table. Through a job called “Requested Nodes,” Records Management notifies Content Move to re-evaluate an item with the next scheduled job.

Content Value versus Content Cost

For every piece of content, there is a cost associated with storage, back-up, and administration. If the chosen storage option does not change over time, storage costs also remain unchanged, even as content ages and loses value. Furthermore,
unmanaged content typically lacks a defined destruction date. In theory, the expense (and risk!) associated with documents and other enterprise content exists forever, unmitigated.

The value, or business benefit of enterprise content declines over time. Whitepapers, market analyses, financial reports and other business records are most valuable at initial creation. As business scenarios change and documents based on new information are published, the worth of existing content declines, until (in many cases) older content no longer supplies useful, reliable data for business decisions.

**Figure 8:**
The value of content is at its highest at publication and declines over time. The top graph shows that the cost of storing and managing that content is high and static without Content Archiving. The bottom graph shows that with Content Archiving, storage costs decline over time as the value of the content declines.
Considering that most business content follows a prescribed lifecycle, using Intelligent Storage Management (ISM) to store, archive, and destroy documents on the basis of business value is certainly valuable. Today, there are various storage options ranging from very fast, very large and very expensive enterprise SAN solutions, to far less expensive and lower performance SATA-based systems. The key problem for most ECM applications is the inability to use available hardware technologies to manage content effectively.

In stark contrast with most disjointed Hierarchical Storage Management (HSM) systems, which infer ideal storage location by reacting to external events (e.g. not accessed for 90 days), Open Text Archiving for Content Server allows the context of content to determine the appropriate storage location. Archiving for Content Server offers automatic, intelligent selection of storage locations, taking advantage of a business record’s defined lifecycle and other uniquely-specified criteria.

ISM is a superset of Hierarchical Storage Management (HSM); it can do everything HSM can do and more. While HSM is essentially always a reactive solution, ISM can be as proactive as it is reactive.