

**Integrating Content Management
Within Enterprise Applications:
*The Open Standards Option***

© Copyright Xythos Software, Inc. 2005
All Rights Reserved



Table of Contents

Introduction	3
Why Developers Are Choosing Open Standards Content Management Platform	3
Approaches to Developing a Content Management Platform	4
Xythos WebFile Server Overview	5
<i>Scalability</i>	6
<i>Integration</i>	7
<i>Data Integrity and Recovery</i>	7
Customer Examples	8
<i>SAS</i>	8
<i>Chubb Insurance</i>	8
<i>Blackboard</i>	8
Conclusion	9

Introduction

In today's computing environment, unstructured data is pervasive. Most business processes require unstructured information, such as legal forms, documents, medical images, presentations, invoices, and diagrams, and involve multiple participants. Frequently, the unstructured data must be shared and managed by participants who reside in different locations. As a result, an increasing number of enterprise applications need to interact with a variety of types of unstructured information and support collaborative document and file management capabilities. This white paper will explore why application developers are increasingly relying upon an open standards Content Management Platform to cost-effectively integrate web-enabled document and file management functionality within enterprise applications.

Why Developers Are Choosing an Open Standards Content Management Platform

As an increasing number of enterprise applications have collaborative requirements, a common infrastructure is needed to manage unstructured information, in much the same way that an RDBMS is used for structured data. Most applications are not designed to cohesively manage unstructured information. As a result, files are exchanged outside of a given application via email and attached to a piece of structured information, such as a sales record in a Customer Relationship Management (CRM) system. While the attachments are helpful, they are static and do not support collaboration among multiple users.

For example, a sales representative may be working on a proposal document and a presentation with several members of his account team. In most systems, he can attach a copy of the document and presentation to the appropriate customer record. However, the CRM system does not allow the team to dynamically edit, change and share the files with each other and centrally manage the information. More often than not, the document will be removed from the system and sent via email, thus keeping the CRM system from containing the most recent version or any history of the document's lifecycle.

Currently, most applications store unstructured information inside relational databases (as BLOBs) or a file system with a reference pointer to the database. The database-centric approach introduces performance overhead and eliminates access and usage of the native file system and its utilities. The database approach also requires development of additional file management services that are not available via the database. Without using a database, some developers have gone to the expense of developing and maintaining application specific library services to facilitate managing unstructured information.

Just as a relational database management system provides the necessary infrastructure to manage structured information, a Content Management Platform is needed to provide an application with the necessary infrastructure to manage and share documents, presentations, spreadsheets, reports, images and other types of files.

Approaches to Developing a Content Management Platform

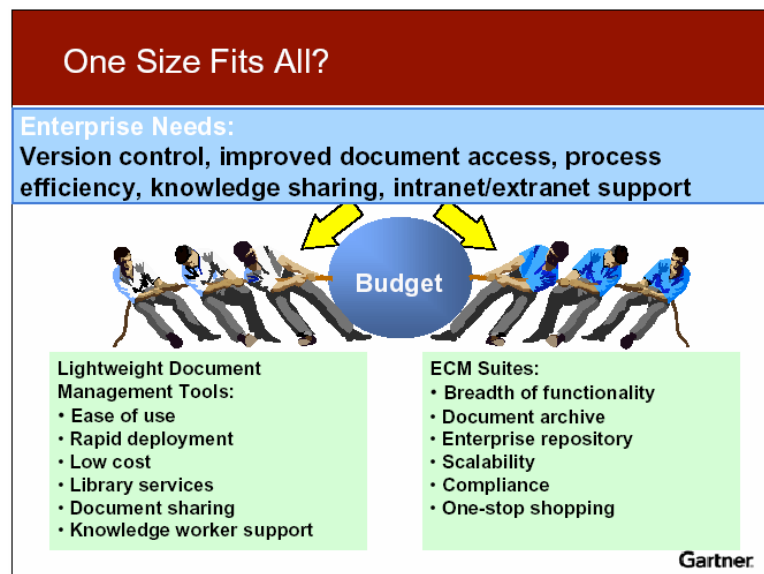
Traditionally, application developers seeking to integrate document or web-enabled content management functionality within their environment consider the following alternatives:

- 1) Building a system by relying on a relational database as the content repository
- 2) Building a system using an open source framework
- 3) Integrating a commercial Enterprise Content Management (ECM) system
- 4) Selecting an open standards-based Content Management Platform

Option 1 may appear to be the least costly alternative, but it often requires the greatest amount of effort. Internal teams or consultants must develop or have the skills needed to address security, reliability, and scalability issues associated with web architecture, in addition to understanding how to combine fundamental content management features within an enterprise application. While these challenges can be overcome with enough time and resources, the results are generally not portable to other applications and require specialized maintenance as other system components change. The time-to-market cost associated with this method is high, particularly where content management is not the core skill set of the development team.

Option 2 is not a viable option for large deployments, where scalability and support are important issues. While several open source projects have emerged to provide unstructured data management, they have limited feature sets, little to no product documentation, and do not provide the scalability and support needed for large deployments. In order to extend the open source system's capabilities, an application vendor must become an active contributor to the open source project, which can introduce substantial cost and project delays.

Option 3 typically involves license fees which make it prohibitive for the purposes of simply integrating document or web-enabled content management functionality within an enterprise application. Traditional ECM vendors such as Documentum and FileNET tend to have complex, legacy product architectures which do not lend themselves well to third-party integration. These products are better suited for managing highly regulated document management processes, such as drug discovery and development, but were not designed to be web-enabled system components for integrating structured and unstructured content within another enterprise application. The



total cost of customization and integration they typically require, often makes them a costly option for content management integration projects. ECM products will certainly play a critical role in managing content within an enterprise, but as Gartner analysts have projected, by 2008, 75% of Global 2000 organizations will have two fully integrated, but very different document management implementations¹. Besides ECM products, organizations will be looking for lightweight document management tools and an underlying platform which will allow developers to web-enable and manage unstructured content within and between a variety of enterprise applications and business processes.

The most effective solution for application developers seeking to integrate document or web-enabled content management functionality within their environment to consider is Option 4, selecting an open, standards-based Content Management Platform like the Xythos WebFile Server (WFS).

Xythos WebFile Server Overview

Xythos WFS provides ubiquitous, high-speed access to documents, simple file management, and library services that can be easily integrated within enterprise applications and presented to end users in a variety of methods. WFS was designed to support large-scale deployments of enterprise applications to tens of thousands of users from a single instance; it also provides a set of useful and immediately deployable services to application developers. Using the WebDAV (Web-based Distributed Authoring and Versioning) protocol and a documented API (Application Programming Interface), developers can integrate a standard set of library services such as version control, file and folder-level security, comments, check-in/check-out, flexible metadata and full file logging to support the collaborative requirements of their applications.

There are four main components to the Xythos WFS product architecture:

- **File System:** WFS stores all file content directly onto a storage device using its native file system. For example, WFS can store information in Networked Attached Storage (NAS), a Storage Area Network (SAN), or in Directly Attached Storage. WFS is also designed to scale its storage capacity by simply adding additional storage devices. There is no limitation to the amount of space and the number of storage devices that WFS can manage; WFS can easily scale to store any size file and any number of files whether they are simple office documents or large image and video files.
- **Database:** Xythos uses a relational database to store all file metadata, such as versions, properties, comments and system logging information. Supported databases include: DB2, Oracle, Postgres (open source), and MS SQL.
- **Server:** Similar to storage capacity, WFS can also scale across multiple databases. Multiple, independent databases can be part of the same WFS installation. Therefore, there is no limitation upon the volume of files and associated metadata that can be managed by WFS.
- **Application Server:** WFS is a java servlet application that runs in any J2EE-compliant application environment. As with storage and metadata scalability, the application server component does not limit the size of the WFS system that can be implemented. Since WFS does not store state information in the application

¹ Karen M. Shegda, Gartner, Gartner Symposium in Orlando, Florida, October 2004

server, the system can be scaled by adding multiple, independent application servers that can be run together in the same WFS installation.

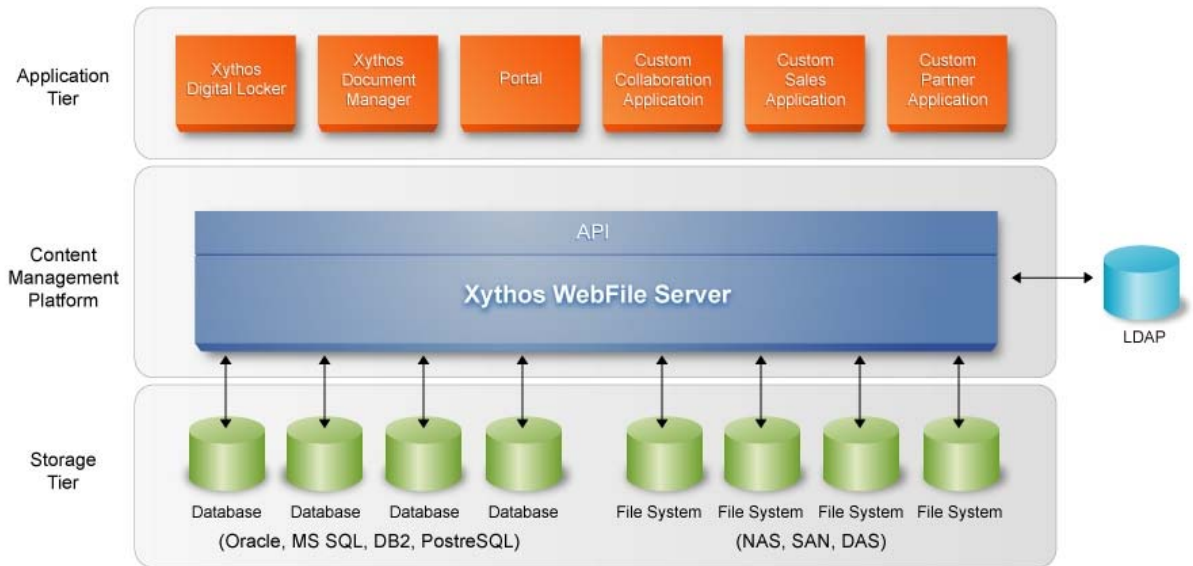


Figure 1: WFS architecture is based on open web-based standards

Scalability

The Xythos WFS was designed to easily scale to meet the most demanding performance requirements. Scalability is achieved through a loosely coupled distributed system in which each component can scale independently. WFS can operate with all of its components on one machine or the components can be deployed on separate machines. Installations with large numbers of users and data can be accommodated by increasing the number of each respective storage, database and application server component. WFS performance has been optimized for HTTP and WebDAV (Web Distributed Authoring and Versioning) so that it can efficiently support file access and storage over long distances; it employs a stateless connection between client and server to support massive user systems requests and overcome connection service disruptions, unlike other LAN protocol-based systems.

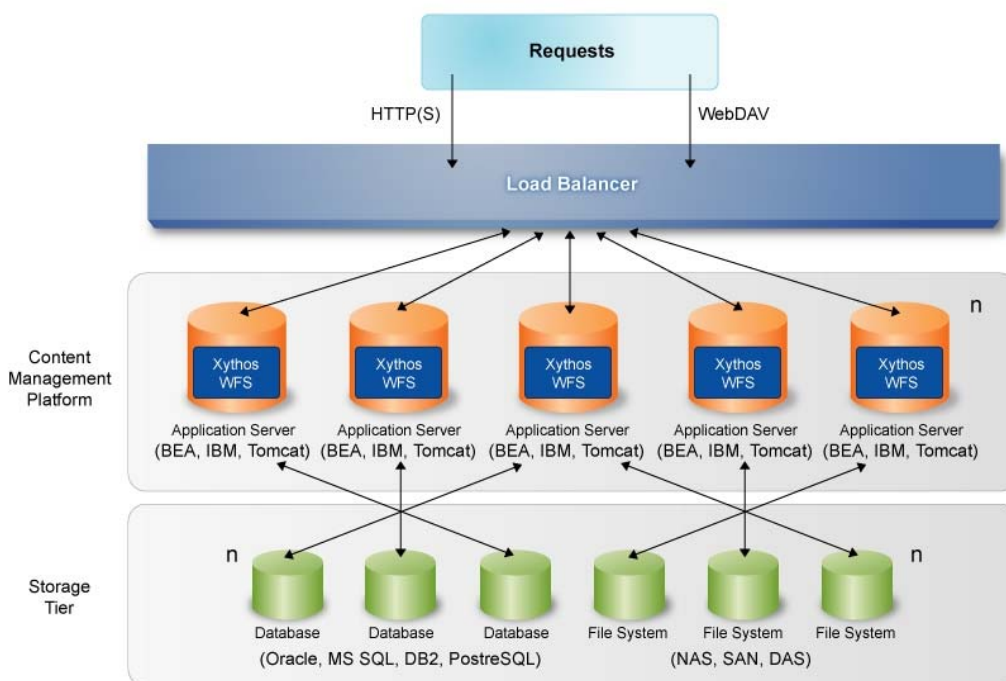


Figure 2: WFS scales with the number of databases, files or users an organization has.

Integration

WFS can be used to add unstructured information management and collaborative capabilities to any enterprise application. The WFS API can be used to integrate all of WFS's services into an application where application logic is coded directly on top of the WFS core feature set. Developers can also easily limit which WFS features their web application presents based upon project and security requirements.

Applications also can communicate directly with the WFS using the WebDAV extension to the HTTP protocol. The first and most widely deployed web service, WebDAV allows client applications to perform remote file management via a standard extension to HTTP. As a result, other applications can communicate with WFS, allowing access, storage and management of active files using a URL that resides in the client application. The capabilities of WebDAV can also be extended, using the WFS API to address file auditing, change status notification, and other security and compliance related requirements.

Data Integrity and Recovery

Data integrity can be ensured by enabling point-in-time recovery for all metadata stored in the relational database and for all file structures residing in external storage (the file system). By adhering to the regular backup schedule of the DBMS, point-in-time recovery of the metadata can be accomplished via standard database utilities that replay transaction logs written since the last backup. By setting up two storage locations (primary and temporary), located on separate disks, WFS can store a backup copy of the actual files. The temporary storage can be configured so that the redundant files are deleted after each successful file backup. Thus, the temporary storage location functions similarly to a database redo log. For the recovery of individual files, WFS provides a trash can and a file access history audit log. The trash can maintains all deleted files until it is emptied



by those authorized to do so. After the trash can has been emptied, an audit log containing a deleted file history can be searched to access and recover deleted files, overcoming this common file recovery limitation associated with database-only dependent systems.

Customer Examples

SAS

SAS is a market leader in providing a new generation of business intelligence software and services that create true enterprise intelligence. SAS integrated the Xythos WebFile Server to enable document management and collaboration within a variety of its enterprise applications. One example is the SAS Drug Development (SDD) application used by life science companies to support the process of developing new pharmaceutical products. SDD uses WFS to manage and share files and documents between researchers, managers and regulators to support the clinical trial reporting requirements during the FDA regulated discover and approval process. SAS Corporate Compliance for Sarbanes-Oxley allows companies to establish a secure, central framework for documenting and reviewing disclosure controls and procedures related to SEC filings. Using WFS, SAS provides a collaborative environment for document preparation, review and certification to support the law's financial reporting requirements. The SAS Enterprise Business Intelligence Server stores business intelligence information and reports in WFS, which allows geographically distributed users and applications to securely access, manage, and share both structured and unstructured business intelligence information.

The Chubb Corporation

Founded in 1882, Chubb is a leading provider of commercial, specialty and personal insurance to customers worldwide and ranks among the top 10 publicly traded insurance organizations based on revenues in the United States. To cost effectively meet advanced business needs for document access, Chubb migrated its original commercial underwriting document repository to the Xythos WFS. The enhanced performance, security and programming features of the WFS have permitted Chubb to provide underwriters with secure access to submissions, correspondence, reports and other documentation using standard web browsers. Using WFS, Chubb now has one central location for managing, tracking and storing documents, diagrams, and reports needed for commercial underwriting, giving the company one system for managing underwriting across the globe.

Blackboard

Blackboard is the leading provider of software and services to the education industry. The Blackboard Content System supports the collaborative processes associated with Learning Content Management, e-Portfolio Management, Virtual Hard Drive Management, and Library Digital Asset Management. The WFS has been integrated with the Blackboard Content System to enable students, faculty and researchers to access, store and manage content associated with the above processes.



Conclusion

A content management platform provides organizations with the flexibility to deliver unstructured data management within the applications they already own and use, reducing implementation and support costs while providing the flexibility required for changing business conditions. Open standards-based content management platforms, such as the Xythos WebFile Server can also provide the core library services functions once only available in high end ECM systems providing application developers with a significantly lower cost alternative with which to address enterprise document and file collaboration requirements. The Xythos WFS has been chosen and deployed to over 200 customers worldwide as a simple document application and content management platform. Software partners have also leveraged the WFS in their product lines to provide content management and enhanced application collaboration capabilities to its customers. For more information on Xythos, go to: www.xythos.com.