Delivering Shared, Authenticated, Ubiquitously Accessible Storage via WebDAV

Michael Bourque, Boston College
Stephen D. Franklin, University of California, Irvine
Daniel A. Updegrove, The University of Texas at Austin
Overview

• Steve Franklin –
  WebDAV
  University of California, Irvine’s Choices

• Mike Bourque –
  Boston College
  Goals & Implementation

• Dan Updegrove –
  The University of Texas at Austin
  Goals & Implementation
WebDAV

- **Web Distributed Authoring and Versioning** (WEBDAV/webdav) Access Control Protocol
- IETF open standard(s), defined by RFCs
- IETF WebDAV Working Group Charter: “Extensions to HTTP that enable remote collaborative authoring of Web resources”
- Collaborative Authoring $\leftrightarrow$ Shared, Authenticated
- HTTP/Web $\leftrightarrow$ Ubiquitously Accessible
- Web Resources $\leftrightarrow$ (Access Controlled) Storage
WebDAV Extensions

- Retrieve and Update Properties
- Access Controls
- Make and Manipulate Collections
- Lock/Unlock
- Versioning
- Atomic Actions

Technical “adjustments” for collaboration actually easier than the necessary Human Social/Mindset adaptations
User Interaction with a WebDAV server

• Via a WebDAV client, special purpose or part of the OS as with Microsoft XP’s WebFolders

• Using a Web Browser connecting to an HTTP/WebDAV Server connecting to a WebDAV Server accessing the resources:
  
  Web Browser ↔ HTTP/WebDAV ↔ WebDAV

• Authentication (and access permissions) via
  – Login (user id and password)
  – Xythos: Resource-specific “ticket” (and password)
Xythos Webfile Server Architecture

Application/Web Server: Apache/Tomcat, IIS, WebLogic, Websphere

<table>
<thead>
<tr>
<th>Web Server Servlet Engines</th>
<th>Custom Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>WFDM</td>
<td>WebDAV</td>
</tr>
<tr>
<td>WebFile Server</td>
<td>Storage Manager</td>
</tr>
<tr>
<td>File Manager</td>
<td>Metadata Manager</td>
</tr>
</tbody>
</table>

Workflow Engine

File System

SQL, Oracle DB2

LDAP, AD, NDS

© Xythos, 2005
University of California, Irvine

2005 marks UC Irvine’s 40th Anniversary

One of 10 campuses of the University of California

Distributed administrative structure

- 19,417 Undergraduates (2003-04 3-term average)
- 4,658 Graduate Students (3,533 Campus, 1,125 Health Sciences)
- 1,400 Teaching Faculty
- 4,610 Campus Staff, 3,480 Medical Center Staff
Anteaters & WebDAV

- Anteater = UC Irvine Mascot
- Anteaters = UC Irvine Alumni
- Paul Mockapetris, Ph.D.: DNS (and SMTP)
- Roy Fielding, Ph.D.: http/1.1, Apache
- Jim Whitehead, Ph.D.: WebDAV
- Professor Richard Taylor: Fielding, Whitehead
UC Irvine Objectives

• Institutional (vs personal) storage
• “Ubiquitous” access
• Alternative to large e-mail attachments
• Sharing/Collaboration
  – Check-out/in, Locking & logging, Version control
• (“Limited”) Web publishing

• Secure Electronic Information Resources
UC Irvine Client Service Base

• UCI Faculty, Grad Students and Staff
  – Substantial quantity (e.g., 1GB/person)
  – UCInetID (Kerberos) Authentication

• Private or Easily Shared
  – With UCI and with non-UCI Collaborators
    Xythos “tickets” essential for the latter
Done Other Ways at UC Irvine

• Course DropBoxes
  – Specialized Application
    • Time-dependent change of access controls
    • Access depends on fine-grain registration status
  – Separate service provided as part of UC Irvine’s EEE = Electronic Educational Environment
    • http://eee.uci.edu/toolbox/dropbox/

• Individual/Personal (Student) File Store
  – Mobile personal memory devices USB connected, based on flash memory or mini-disks (e.g., iPod)
Xythos at Boston College

Michael Bourque
Associate Vice President
Information Technology
Services

September 21, 2005
Agenda

• About Boston College
• Web Storage Requirements
• Xythos Deployment at BC
• System Topology
• BC Custom Changes
• User Examples
• Next Steps & Issues
About Boston College

• Private, Catholic University in Chestnut Hill, MA
• Founded by the Jesuits in 1863
  – 9,060 undergraduates
  – 4,755 graduate students
  – 640 Full-time faculty
  – 2,310 Staff
• Relatively centralized University Administration & ITS
University Web Storage Requirements

- Provide a secure environment for sensitive data
- Allow access to the files from any location
- Allow easy access to files from multi-platform clients (e.g. via WebDAV)
- Make it easy to share files with BC and non-BC people alike for collaboration purposes
- Include versioning, locking and logging capabilities
- Give users the ability to control access – both individual and role-based
- Provide an alternative to emailing files
- Provide a way to back-up important documents
Xythos Deployment at BC

• Planned Slow Rollout
  – Sep 2003 - Beta release to Tech folks & pilot groups
  – Jan 2004 – Faculty & Staff Release 1GB quota each
  – Jun 2004 - Student Release 100MB quota each

• Usage
  – Users – 17,000
    • Faculty/Staff – 2,200
    • Students – 11,200
    • Other - 3600
  – Number of files – 1.1 million
  – Total space used - 270 GB
    • Faculty/Staff – 132 GB
    • Students –60 GB
    • Other – 78 GB
  – Total usable space – 5.5 TB
System Topology

Clients

Web Folders
Web Browsers

F5 BigIP pair

App Servers

App Server 1
AIX 5.2
Tomcat 5.2.8

App Server 2
AIX 5.2
Tomcat 5.2.8

Data

Netapp Filer
6.5 terabytes
protocol: nfs

Oracle database 9.2.0.4
AIX 5.2

Data

File metadata

Application data

Shark

CVS Server

LDAP Servers
BC Custom Changes

• Significant user interface customizations
  – These include integration with WebMail, CampusGroups and Directory (LDAP) Search

• LDAP Integration with our custom BC schema for authentication, file access/permissions and account creations
  – This includes a custom group implementation that relies solely on LDAP for associating users to groups.

• Help Documents

• Ability to implement changes without system downtime
### BC Custom Changes - User Interface

#### MyFiles@bc

**Velisse Kuran**
- **home** (Volzi)
- my bookmarks
- my searches
- contacts
- options
- trash - empty now
- sign out
- help & support

#### MyFiles@bc Interface

1. **Folder:** A folder named `adiazi`. The folder contains various files and directories.
2. **File Structure:**
   - `PasswordProject`
   - `CampusGroups`
   - `WebMailProject`
   - `XythosProject`
   - `VoicemailProject`
   - `trash`
   - `Testing`
   - `Printers`
   - `Eve-Stuff`
   - `EPO`
   - `RromenCommittee`
   - `DoxExpress`
   - `ProjectOffice`

#### File Details:
- **Name:** Various files like `PasswordProject`, `CampusGroups`, etc.
- **Size:** Various sizes, e.g., `3.26M`, `456K`, `33.12M`, etc.
- **Date Modified:** Dates range from `5/21/04` to `9/28/04`.
- **Permissions:** Icons indicate read, write, execute permissions.

#### Additional Information:
- **Quotes:** Appears to be a quota related to file storage.
- **System Information:** Shows the status of the system.

---

**WebDAV: Bourque, Franklin, Updegrove**
BC Custom Changes – LDAP Integration

Find Users and Groups

Select your chosen user(s)
- Display Name
- Title
- Anderson, Adrian
- Anderson, Alberta P
- Anderson, Alexis J
- Anderson, Allison A
- Anderson, Andrea A
- Anderson, Annemarie P
- Anderson, Bonnie
- Anderson, Courtney G
- Anderson, Danielle M
- Anderson, Denise N
- Anderson, Emily A
- Student in College of Arts & Sciences
- Assoc. Clinical Professor of Language
- Student in College of Arts & Sciences
- Student in College of Arts & Sciences
- Student in Graduate School of Arts
- Student in School of Nursing, class of 2008
- Student in Graduate School of Nursing

MyFiles@bc

Nelisse Kuran
- home (/d/az)
- my bookmarks
- my searches
- contacts
- options
- trash - empty now
- sign out
- help & support

Find Users and Groups

Select your chosen user(s)
- Display Name
- Academic Technology Services
- Information Technology Services, Admin Services
- Information Technology Services, Appl & Syst Svcs
- Information Technology Services, Application Svcs
- Information Technology Services, Computing Support
- Information Technology Services, Network Services
- Information Technology Services, Office of the VP
- Information Technology Services, Policy & Security
- Information Technology Services, Prod & Data Svcs
- Information Technology Services, Project Mgmt Off
- Information Technology Services, SLSC
- Information Technology Services, Training & Commun
BC Custom Changes – Help Documents

MyFiles@bc

Getting started with MyFiles@bc
- What is MyFiles@bc?
- Touring the MyFiles@bc Web application
- Getting started using Windows
- Getting started using Macintosh
- Frequently asked questions

Learning to use MyFiles@bc
- Managing your account
- Managing your files
- Managing your folders/directories
- Sharing your files and folders
- Searching
- Bookmarking

New to MyFiles@bc?
- Why use MyFiles@bc? (PDF)
- Learn how to use MyFiles@bc for your classes with the MyFiles@bc Professor Reference (PDF)

Search Help Center

Worried about Viruses?
MyFiles@bc does not scan files for viruses at upload or download. The best way to protect yourself is by updating your anti-virus software and scanning your office and home computers routinely. Learn more...

Access MyFiles@bc on the Web:
www.bc.edu/myfiles
User Examples

• Project Collaboration
  – Residential Life Staff Secure Intranet
  – University Planning & Assessment Committees
  – Information Technology Services
    • Project Management Office
    • Systems & Mgmt Operation

• Sharing Large Files
  – Research Professor – Uploading recorded interviews
  – Marketing Office – Sending video files to offsite vendors for processing

• Group folders for departments and student organizations
  – Sensitive data storage

• Alternative to emailing file to yourself
Next Steps and issues

• Next Steps
  – October - upgrading to Xythos WFS 4.2.34-2
  – Implementing account purge process
  – Continue Training Program and Lunch & Learns

• Issues/Challenges:
  – Apple’s WebDAV implementation is significantly slower than Windows
  – Some issues with web folders freezing with MS Office 2003
  – Would like to have a multi-platform desktop backup solution
  – Still ironing out how to handle quota increase requests for students
  – Would like to integrate with email (substitute for local folders)
The University of Texas at Austin

- Founded in 1883
- Flagship of the 15-campus University of Texas System, with 6 medical centers
- 51,000 students; 11,000 degrees/year
- 300,000 continuing education enrollments
- 3,000 faculty, 18,000 staff
- Over 450,000 alumni
Overview

- WebSpace/Xythos at UT Austin
- Current Stats
- Issues
- Part of an Evolving Trend
WebSpace/Xythos at UT Austin

• Large scale ubiquitous web storage – local implementation of Xythos is called “WebSpace”
• File storage and web publishing system provided to all students, faculty and staff at UT
• As part of the University’s instructional technology fee, each student will receive a “free” storage allocation, which can be used to store files and publish personal Web pages – the new “floppy”
## Current Stats

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Student</th>
<th>Staff</th>
<th>Faculty</th>
<th>Alumni</th>
<th>Member</th>
<th>Affiliate</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total in ADAM</strong></td>
<td>224514</td>
<td>50695</td>
<td>10264</td>
<td>3625</td>
<td>42049</td>
<td>91</td>
<td>36470</td>
<td>81320</td>
</tr>
<tr>
<td><strong>WebSpace Account</strong></td>
<td>61596</td>
<td>33740</td>
<td>3388</td>
<td>1306</td>
<td>17128</td>
<td>6</td>
<td>3945</td>
<td>2083</td>
</tr>
<tr>
<td><strong>UTSystem</strong></td>
<td>330</td>
<td>19</td>
<td>1</td>
<td>8</td>
<td>71</td>
<td>1</td>
<td>16</td>
<td>214</td>
</tr>
</tbody>
</table>

**Student** - student-current
**Staff** - staff-current, staff-budgeted
**Faculty** - faculty-current
**Alumni** - graduate

**member** - faculty-future, staff-future, official-visitor, visiting-scholar
**affiliate** - prospective-student, student-admitted, extension-student, retiree, donor-friend-vip, library-patron, business-entity
**None** - former-student
Current Stats

- Total File Space Used GB: 946.38
- File Space Available GB: 644.78
- Avg HOME Directory Size MB: 17.01
- Average data transferred per day: 56.37 GB
- Around 82% of the requests that come to WebSpace, come from Windows based computers
- Around 39% of the off campus traffic comes from Road Runner
- 75-90% of the webspace traffic comes from off campus addresses
- 1.5-2% of the total WebSpace requests come from WebDAV clients
Evolving Trend

• “Proponents of the Web 2.0 approach believe that web usage is increasingly oriented toward interaction and rudimentary social networks, which can serve content that exploits network effects with or without creating a visual, interactive web page. In one view, Web 2.0 sites act more as points of presence, or user-dependent web portals, than as traditional websites.”

• Web 2.0 and WebSpace – no home page, portions of content shared out of this account with varying levels of access controlled by user
Issues

• Lacking real data on how people are using it
  – Bandwidth limit, storage limit experiment
  – Determine what users expect from WebSpace

• Need for large scale admin tools
  – Bandwidth management
  – Storage limit management

• Support issues
Contact/Reference Information

• Michael Bourque <mike@bc.edu>
  Boston College
  MyFiles: www.bc.edu/myfiles
• Stephen D. Franklin <franklin@uci.edu>
  University of California, Irvine
  Webfiles: webfiles.uci.edu
• Daniel A. Updegrove <updegrove@mail.utexas.edu>
  The University of Texas at Austin
  WebSpace: webspace.utexas.edu
• WebDAV: http://en.wikipedia.org/wiki/Webdav
• Higher Ed Webfile (Xythos WebDAV) mailing list:
  https://maillists.uci.edu/mailman/listinfo/webfile-hied