

WebDAV and Apache

Greg Stein

gstein@collab.net

<http://www.lyra.org/greg/>

Agenda

- Overview
- Benefits
- How does it work?
- Some scenarios
- DAV software
- Setting up mod_dav
- Futures

What is WebDAV?

(1 of 2)

- **Web-based Distributed Authoring and Versioning**
 - “DAV” is the usual short form
- Goal: enable interoperability of tools for distributed web authoring
- Turns the Web into a *writable* medium

What is WebDAV?

(2 of 2)

- Applies to all kinds of content - not just HTML and images
- Based on extensions to HTTP
- Uses XML for properties, control, status
- RFC 2518

Benefits

- Benefits for all web users:
 - Users
 - Authors
 - Server administrators
- Technical benefits for developers, network administrators, and security personnel

User Benefits

- User: defined here as a web surfer
- Document metadata available
- More intelligent “directory” listings

Author Benefits

- Author: the person who writes the content
- Standard way to place content on server
- Move/copy the content around
- Tag the content with metadata
- Overwrite protection in group scenarios

Administrator Benefits

- Administrator: the person running the server
- All interaction via the protocol
- Divorces local system layout, config, and structure from the author's conceptual space
- HTTP-based authentication instead of system accounts

Technical Benefits

Overview

- Properties (“metadata”)
- Overwrite protection
- Namespace management
- Versioning
- Infrastructure: old and new
- Replacement protocol

Technical Benefits

Terminology

- Collection
 - A collection of resources
 - A collection is also a resource
- Resource
 - Generic name for collections or *member resources*
- Member Resource
 - “Leaves” in a URL namespace

Technical Benefits

Properties

- Properties are name/value pairs
 - Names are uniquely identified with URIs
 - Values are well-formed XML fragments
- All resources have properties
 - Files and directories
 - Server-defined/maintained, or client-defined
- Records metadata such as author, title, modification time, or size

Technical Benefits

Overwrite Protection

- Shared and exclusive locks
- Locks have characteristics such as timeouts, owners, and depth
- Identified by authentication and lock token
- Apply to whole resources, not portions

Technical Benefits

Namespace Management

- “Namespace” refers to the URL hierarchy
- DAV provides mechanisms to create, move, copy, and delete resources

Technical Benefits

Versioning

- Woah... big topic
- “DeltaV” – RFC 3253
- Simple, linear versioning, or complex configuration management
- Client-side and server-side workspaces
- “Baselines” are snapshots
- “Activities” can act as change sets

Technical Benefits

Existing Infrastructure

- Receives benefits of HTTP infrastructure
 - Strong authentication
 - Encryption
 - Proxy/firewall navigation
 - Worldwide deployment
 - Huge talent pool; numerous tools, apps, etc
- More on this later

Technical Benefits

New Infrastructure

- DAV can provide infrastructure for:
 - Collaboration
 - Metadata
 - Namespace management
 - Versioning
 - *Ordered collections*
 - *Access control*
 - *Searching*

Technical Benefits

Replacement Protocol

- DAV providers read/write to the web server
- Can obsolete other mechanisms:
 - FTP
 - FrontPage and Fusion proprietary protocols
 - Custom or one-off solutions
- Robust enough for future enhancements

How Does it Work?

- A protocol layered on HTTP/1.1
 - HTTP/1.1 clarifies the extension process
- HTTP extensions
 - New HTTP headers
 - New HTTP methods
 - Additional semantics for existing methods

New HTTP Headers

- DAV:
- If:
- Depth:
- Overwrite:
- Destination:
- Lock-Token:
- Timeout:
- Status-URI:

New HTTP Methods

Overview

- COPY, MOVE
 - MKCOL
 - PROPPATCH, PROPFIND
 - LOCK, UNLOCK
-
- Eleven new methods for DeltaV

New HTTP Methods

COPY, MOVE

- Pretty obvious: copy or move resources
- Copying collections uses Depth: header
- Destination: header specifies target
- Also uses Overwrite: header
- Optional request body controls the handling of live properties

New HTTP Methods

MKCOL

- Create a new collection
- Avoids overloading PUT method

New HTTP Methods

PROPPATCH, PROPFIND

- PROPPATCH is used to set, change, or delete properties on a single resource
- PROPFIND fetches one or more properties for one or more resources

More on PROPFIND

- Using PROPFIND anonymously allows users to discover files
- Best to require authentication
- In the future:
 - Browsers will want it for “nice” directories
 - Clients will want PROPFIND for metadata
 - Server will have finer granularity to hide items

New HTTP Methods

LOCK, UNLOCK

- Add and remove locks on resources
- Both use the Lock-Token: header

Futures: WebDAV

- Access Control (submitted; Q4 2002?)
- Advanced Collections
 - Bindings (restarting)
 - Ordering (idle)
 - References (idle)
- Searching (progressing; Q2 2003?)

Scenarios

- Departmental Server
- Web Hosting
- Software development teams
- Remote collaboration
- Network file system
- Unified repository-access protocol
- Application protocol

Scenario: Departmental Server

(1 of 2)

- Department of 20 staff
- They operate a private web server
- Web server acts as a repository
 - File servers used to play this role
- Everybody needs to author documents
- Web server (vs file server) provides better navigation, overviews, and offsite links

Scenario: Departmental Server

(2 of 2)

- Web site is DAV-enabled
 - Allows remote authoring and maintenance
 - Allows tagging documents with metadata
- Security can be used to limit or partition areas for specific users
- Documents drop right onto the server
- New pages for summaries and overviews

Scenario: Web Hosting

(1 of 2)

- 5000 users
- `http://www.someisp.com/username/`
- No need to enter users into `/etc/passwd`
 - Use any Apache `mod_auth_*` module
- User directories can be distributed, shifted, updated as needed across the filesystem

Scenario: Web Hosting

(2 of 2)

- Apache's httpd.conf gets complicated
 - Need section for each user
 - Something like UserDir would be great
 - For now, include a generated file

WebDAV Software

Clients

- Joe Orton: cadaver, sitecopy, Neon
 - Nautilus, GNOME, KDE, Goliath
 - SkunkDAV, DAVExplorer
 - APIs: Python, Perl, C, Java
-
- Commercial: Microsoft, Adobe, Macromedia

WebDAV Software

Servers

- Apache 2.0, and Apache 1.3/mod_dav
- Zope
- Magi
- Tomcat, Jakarta Slide(?)

- Commercial: many

WebDAV Software Systems

- Subversion
- Microsoft Outlook/Exchange

WebDAV Software

Joe Orton's cadaver

- Interactive command-line tool
- Provides listing, moving, copying, and deleting of resources on the server
- Manages properties
- Can lock and unlock resources

WebDAV Software

Joe Orton's sitecopy

- Edit web site locally
- Update remote web site
- Operates via FTP or WebDAV
 - More/better functionality via WebDAV
- Does not do two-way synchronization

WebDAV Software

Nautilus

- Nautilus is the file manager for GNOME
- Uses gnome-vfs
 - “Virtual File System”
 - Can target WebDAV repositories
- GUI-based management of a DAV server

- KDE is DAV-enabled, too

WebDAV Software

Goliath

- Goliath is a DAV client for classic MacOS
- Finder-like
 - Drag and drop
 - Browsing
- Manages locks and properties

WebDAV Software

SkunkDAV and DAVExplorer

- Java “explorer style” WebDAV clients
- SkunkDAV supports content editing
- Both support properties and locks

- SkunkDAV provides a separable library

WebDAV Software

Language APIs

- Good for experimenting and building apps
- Most are layered onto existing HTTP APIs

- Python API from Greg Stein
- Perl API from Patrick Collins
- C API (Neon) from Joe Orton
- Java APIs from SkunkDAV or Jakarta Slide

WebDAV Software

Internet Explorer 5.0

- Enabled with the “Web Folders” add-on
- Adds “Web Folders” section into Windows Explorer, under “My Computer”
 - Allows drag and drop of files
 - Standard move/copy/delete/rename of files

WebDAV Software

Microsoft Office 2000

- Broad distribution
- Word, Excel, etc are DAV-enabled
 - Open/save files directly from/to web server
 - Uses DAV locks for overwrite protection
- First round of Microsoft's move to DAV
 - Also: IIS5, Exchange 2000

WebDAV Software

Adobe GoLive 5.0

- One of the first Web authoring tools to support the DAV protocol
- Page design, authoring, construction
- Uses locking to assist authoring teams
- Site management

WebDAV Software

Apache and mod_dav

- mod_dav provides the DAV support
- Installed on about 250k (public) sites
- De facto reference implementation
 - Class 1 and class 2
 - Extensions for versioning
 - Experimental code for binding, DASL

WebDAV Software

Zope and Tomcat

- Both are application servers
 - Zope is written in Python
 - Tomcat is written in Java
- Zope uses WebDAV to manage content
- Tomcat makes it available, but a good deal of coding is required

WebDAV Software

Subversion

- Open Source version control system
 - Intended to replace CVS
 - Fixes CVS problems, adds improvements
- Subset of DeltaV for its network protocol
- Lots of leverage: Apache 2.0, Berkeley DB
- Reusable libraries

Setting up Apache/mod_dav

Overview

- Grab and install tarball
- One simple directive:
DAV On
 - Use within <Directory> or <Location>
- Need to change file/dir ownership and privs
- Enable locking
- Add security as appropriate

Basic Installation

- Grab tarball
 - <http://www.apache.org/dist/httpd/>
- Pass `--enable-dav` and `--enable-dav-fs` to the `./configure` script
- May also want `--enable-auth-digest`

Example Configuration

```
Alias /gstein /home/apache/davdirs/gstein  
<Location /gstein>  
    DAV On  
</Location>
```

Filesystem Changes

- Assume Apache is run with UID “nobody” and GID “www”

```
% ls -la /home/apache/davdirs/gstein
total 3
drwxr-s---  3 nobody   www   1024 Jun 25 14:32 .
drwxr-s---  3 nobody   www   1024 Jun 28 17:26 ..
-rw-r--r--  1 nobody   www    424 Jun 26 16:36 index.html
drwxr-s---  4 nobody   www   1024 Jun 26 13:05 specs
```

Enable Locking

- Additional directive for the lock database
`DAVLockDB /home/apache/davdirs/lock.db`
- Lock databases are per-server

Security Considerations

- Disable bad operations (CGI, includes, etc)
`Options None`
- Prevent `.htaccess`
`AllowOverride None`
- Limit the users' method access
`<LimitExcept OPTIONS GET POST REPORT>`

Limiting PROPFIND

- Note that PROPFIND is in the <Limit> directive
 - Limits the use of PROPFIND to authorized users
 - Based on concerns mentioned earlier about “discoverability” of a web site

Example Configuration

```
<Location />  
  AllowOverride None  
  Options None  
  DAV On  
  AuthName "my web site"  
  AuthType basic  
  Auth_MySQL on  
  Auth_MySQL http_auth  
  <Limit PUT DELETE PROPFIND PROPPATCH MKCOL COPY \\  
    MOVE LOCK UNLOCK>  
    Require user gstein  
  </Limit>  
</Location>
```

Implementing mod_dav

- Apache has great extensibility
- But:
 - Hard to add new methods
 - Security: file ownership, SUID helpers, etc
 - Alternate access to repository
- Security issues led to private repository
- Module provides excellent speed

Futures: mod_dav

- mod_dav 1.0 was released on June 13, 2000
- Apache 2.0 includes core DAV features
 - fully integrated
 - better plug-in system
 - updated, complete versioning hooks
- Apache 2.1
 - Other DAV extensions

Review

- WebDAV can change the very nature of how people interact with the Web
- Great standard, replaces many protocols with a single protocol
- `mod_dav` brings DAV to Apache
- Tools and apps are common and more appearing every day

Resources

- <http://www.webdav.org/>

Everything you need is on this web site, or linked from it.

Q&A

November 21, 2002

ApacheCon US 2002

59