

# CatacombProposal

## Catacomb Proposal

### Abstract

Catacomb is a WebDAV repository module for use with the Apache WebDAV module, `mod_dav`.

### Proposal

Apache `mod_dav` parses WebDAV and DeltaV protocol requests into operations on a repository providing persistent storage of resources and their properties. The default repository for `mod_dav` is provided by a separate module, `mod_dav_fs`, which stores resource bodies as files in the filesystem, and stores properties in a (G)DBM database.

Catacomb provides a replacement for `mod_dav_fs` called `mod_dav_repos` that stores resources and their properties in a relational database using `mod_dbd` from the apache project for database abstraction. The primary advantage of this approach is the searching capabilities of the database are used to implement the DASL protocol. Additionally, the database allows straightforward implementation of the versioning capabilities of the DeltaV protocol.

By shifting to relational database technology, Catacomb is a platform that contains important aspects of typical document management systems: the ability to store large numbers of documents, and search over their metadata. Furthermore, it is possible (via source code modification) to change the set of predefined properties stored in the main schema of the relational database. Properties in the main schema are faster to search.

### Background

Catacomb was initially developed at the University of California, Santa Cruz (UCSC) in 2002 headed by Prof. E. James Whitehead. Catacomb was designed as a reference implementations for WebDAV search and locating DASL. Also, main pieces of the WebDAV-DeltaV protocol (RFC 3253) support originated from this time.

Since 2006 major parts of the development come from the german aerospace center (DLR) and Catacomb gained support for the access control protocol (RFC 3744) and database abstraction with apache `mod_dbd`.

### Rationale

The maintainers and developers of Catacomb are interested in joining the Apache Software Foundation for several reasons:

- Apache has a widely recognized name, which will help Catacomb get an increased visibility and acceptance.
- It might open the door for sharing ideas or cooperation with other projects, such as `mod_dav` or Jackrabbit.
- Catacomb would like to benefit from Apache's infrastructure.

### Initial Goals

Though the bulk of catacombs initial development is complete and the server is running stable, there are still some large areas for future development. Some areas we hope to focus on in Apache:

- Finish of the Access Control Protocol extension
- Building up an automated testing environment
- Implementation of WebDAV transaction methods

### Current Status

#### Meritocracy

The initial development was done at UCSC and released as Open Source software under the Apache Software License. Since release, many developers have adopted the server and submitted significant patches. Later, the DLR take care of the most modifications and changed the license in agreement with all developers to ASL2. Large portions of the codebase are now managed by those most familiar with and responsible for them. Any potentially controversial change is discussed on the public mailing list (<http://catacomb.tigris.org/servlets/ProjectMailingListList>) and good suggestions are frequently implemented.

#### Community

Catacomb is used in many organizations which are interested in the advancement of the catacomb development. Many of these have at least one developer that joined the catacomb mailing list and so the mailing list is the most important communication platform. The catacomb community encourages suggestions and contributions from any potential user and developer.

#### Core Developers

The initial core developers will come from UCSC. Many of these developers are still interested in the advancement of the project. Furthermore there are a lot more committers from other companies which supports catacomb for years now. There is at least one Apache Member together with one other Apache committer.

## Alignment

Catacomb is a module for Apaches Webserver. This is why Apache httpd is the most important dependency for catacomb. And catacomb is also a particularly good fit for Apache due to integration potential with other projects, specifically mod\_dav and Jackrabbit.

## Known Risks

### Orphaned products

Despite to its small number of committers, there is no risk of being orphaned. The project now exists for several years and there is a constant core-community which have a long-term interest in use and maintenance of the code.

### Inexperience with Open Source

Catacomb was started as an open source project in 2002. Many of the committers have experience working on open source projects and there are also at least one developer which has experience as committer on other Apache projects.

### Homogenous Developers

As mentiont above, the current list of committers includes developers from at least two different companies plus many independent volunteers.

### Reliance on Salaried Developers

At this time, many of the code comes from different companies like the DLR. Because the DLR is a research facility, many of the work is done by students working on their diploma thesis.

### Relationships with Other Apache Products

At this time, the only dependency to other apache projects is apaches http server. Another project which implements advanced parts of the WebDAV protocol is Slide/Jackrabbit. But this two project have different objectives and a different target audience so that there is no direct competition between these project. In contrast to that, the developers could learn much from each other when implementing advanced protocol issues.

### A Excessive Fascination with the Apache Brand

The catacomb project exist quite successful on their own and could continue on that path with no problems at all. We expect the Apache brand could help to increase the visibility of the project and so maybe more developers could be interested in the project.

## Documentation

- The existing project page could be found here: <http://catacomb.tigris.org>
- The Catacomb Architecture: [http://catacomb.tigris.org/catacomb\\_arch.pdf](http://catacomb.tigris.org/catacomb_arch.pdf)
- The Catacomb mailing list with archive: <http://catacomb.tigris.org/servlets/ProjectMailingListList>
- The Old mailing list archive: <http://catacomb.tigris.org/oldarchive/threads.html>

## Initial Source

The initial source was hosted at webdav.org. Since 2006 the project is hosted at tigris. The svn repository could be found here: <http://catacomb.tigris.org/source/browse/catacomb>.

## Source and Intellectual Property Submission Plan

The complete catacomb code is under apache software license 2. The complete codebase that is hosted at Tigris will be contributed.

## External Dependencies

At this time, there are no external dependencies.

## Cryptography

None

## Required Resources

### Mailing Lists

- catacomb-dev
- catacomb-users
- catacomb-commits

### Subversion Directory

- per usual ASF guidelines

### Issue Tracking

- Bugzilla (??)

### Initial Committers

- Jim Whitehead <ejw at cse.ucsc.edu> UCSC
- Sung Kim <hunkim at cse.ucsc.edu> UCSC
- Pan Kai <pankai at cse.ucsc.edu> UCSC
- Markus Litz <markus.litz at dlr.de> DLR
- Steven Mohr <steven.mohr at dlr.de> DLR
- Tim Olsen <tolsen718 at gmail.com> LimeWire
- Elias Sinderson <elias at cse.ucsc.edu> UCSC
- Chris Knight <Christopher.D.Knight at nasa.gov> NASA
- Gianugo Rabellino <gianugo at apache.org> Apache
- Ricardo Rocha <ricardo at apache.org> Apache

## Sponsors

### Champion

- Gianugo Rabellino <gianugo at apache.org> Apache

### Nominated Mentors

- Gianugo Rabellino <gianugo at apache.org> Apache
- Justin Erenkrantz <justin.erenkrantz at gmail.com>