



# A new Document Control System

---

“A new system to manage LIGO documents”

Stuart Anderson  
Melody Araya  
David Shoemaker

29 September, 2008

<http://dcc.ligo.org/...>



# The Goal

---

“Setup a document repository that is easy enough to use that it will be widely and regularly used within LIGO (Lab, AdvLIGO, and LSC)”

- Persistent access to documents via the web.
- Self service for adding new content.
- Ability to search for documents by content.
- Compatibility with existing document IDs.
- Access controls.
- Clear separation of public and private documents.



# The Solution

---

“The Giaime commission selected DocDB”

- Open Source tool used at FermiLab (caveat emptor)
- Definition of a document includes multiple files.
- Simple workflows supported, e.g., Engineering sign-off.
- Hierarchy of access controls (group membership).
- Trigger generation (email on document modification).
- Simple version numbering (not version control system).
- Content stored in easy to find files using LIGO document names, e.g., /home/dcc/...T070059.pdf



# Status

---

- Beta system up and running as [dcc.ligo.org](http://dcc.ligo.org).
- Very positive feedback from early testers.
- Few weeks of development remaining for initial deployment.
- Development and support team (Melody) in place.
- Policy team (David and co) in place.



# Initial Phase Tasks

---

1. LIGO document numbers (90%)
2. Public document separation/delay (80%)
3. QA sign off (including dedicated search) (25%)
4. LIGO group hierarchy and topics (defined/refine)
5. Universal search (old and new DCC)
6. S document type for AdvLIGO serial numbers
7. Seed doc numbers (freeze old system before CY09)
8. Users manual (wiki for easy feedback?)
9. [dcc.ligo.org](http://dcc.ligo.org) public redirection and google registration



# Initial Phase Tasks (2)

---

1. Bug tracking system
2. Mailing list
3. Add another digit to document numbers (plan for success)
4. Utilize initial AuthProject tools (Kerberos):
  - No need for another set of shared passwords.
  - No need for users to remember what DCC group to login as.
  - Unique user registration/tracking (who actually added/modified).
  - Automatic setting of Submitter field given unique login
  - Always allow submitter to access file even if not group authorized.
  - Enable richer set of groups in the future.
  - Possibly restrict access to QA field, if really needed?



# Roll Out Plan

---

Ready for phased roll out starting in late October:

- » Phase 1: few weeks of mixed development and beta testers including new testers that were not involved in the initial development (all files are guaranteed to be lost).
- » Phase 2: one month AdvLIGO test (all users required to keep their own local copies until acceptance tests passed).
- » Phase 3: one month of LIGO Lab use (recommend new users keep local copies until they have personally verified they can share and retrieve documents).
- » Phase 4: open to full LSC use sometime in December.



# Near Term Enhancements

---

1. Request specific document ID (e.g., adding old doc).
2. Request specific version number?
3. Automatically increment doc number for the first document of each type in the year.
4. Disallow changing document type when updating version since this is now part of the doc ID and file name.
5. Test migrating old documents.
6. Upgrade backend search engine and possibly add additional frontend.
7. Rebuild and test the Test/Backup system.
8. Test at larger scale.





# Long Term Enhancements

---

1. Update to newer (last?) version of DocDB.
2. Investigate possible uses of event support, e.g., integrate with LSC meeting wiki use?
3. Investigate multi-site replicas (read-only or read-write?) e.g., Observatory access after a backhoe event.
4. Develop, test and document multiple-system failure scenarios.
5. Refactor code based on current experience with changing from native DocDB to LIGO document IDs.
6. Decide whether to fork or attempt to merge back our changes to the DocDB community?



# Open Questions

---

- Should LVC be a separate access group than LSC?
- How fine grained access controls should be supported? E.g., in the long term should the DCC allow working groups to manage private documents, even temporarily?
- Will we be ready to completely freeze the current DCC by December 31, 2008?
- Should we migrate all of the old documents to the new system, if so how long will that take?
- We have invested significantly in a new DCC. What metric should be used to evaluate this systems success?
- How do we do a better job of managing software projects from perceived need, requirements, build vs buy, deployment, support, ...
- How to handle publication documents that we do not own?



# Summary

---

- Beta system is looking very good--end users have indicated that it is a significant improvement over the current system.
- Melody has done a great job on a project with loose and changing requirements.
- Interested ExComm members should upload some test documents. Note, these files will disappear and/or be edited by other testers.