



DocDB Future Directions

Eric W. Vaandering
(ewv@fnal.gov)

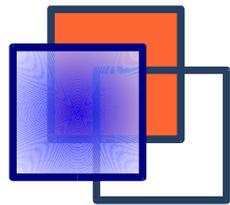
Vanderbilt University

DocDB Stakeholders' Meeting
Fermilab, 31 January 2006

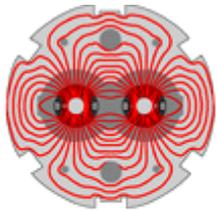


DocDB History

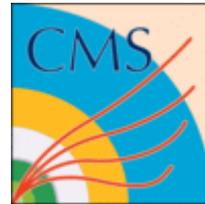
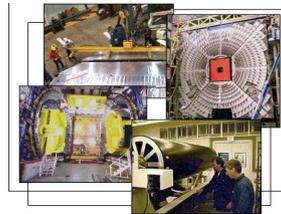
- First demo of protoDocDB: Nov 2001: “What's missing is making the HTML pretty, being able to update a document, multiple file updates, security, and a search and display function.”
- V2 – Keywords, Search: Jan 2002
- V3 - Email Notification: Jun 2002
- V4 - Meeting Organizer: Aug 2003
- V5 - Signoffs, Notes, Meeting Security: Apr 2004
- V6 - Certificates, CSS: Jan 2005
- V7 – Aug 2005: Better events, calendar, XREFs
- Dec 2002: MI (later AD) adopts, Harry Cheung: “Still you are talking about upgrades....”
- Admin interface: Feb 2003
- Mar 2003: DocDB shown at CHEP
- Jan 2004: Gretina adopts, GPL chosen
- Mar 2004: Computing Division adopts
- Jun 2004: CD Contracts for certificates and other improvements
- BTeV Canceled: 7 Feb 2005
- MINOS, Minerva, EPP, LArTPC, and more begin using DocDB



DocDB Usage



LARP



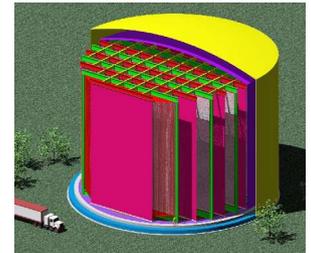
Pixels
Simulations



- Exact numbers unknown, but...



- About 20 groups using DocDB
- Over 1000 authors/users
- At least 10,000 documents managed



LArTPC



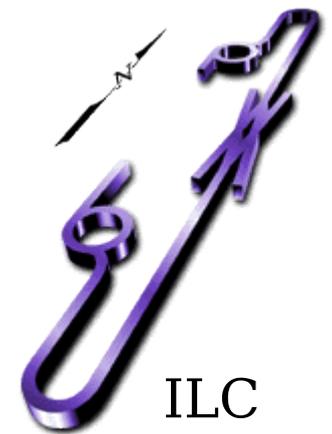
OSG



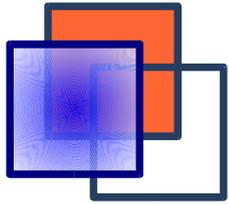
**A SEARCH FOR
NEW PHYSICS**
weak



DARK ENERGY
Survey

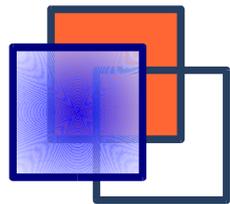


ILC



Some Numbers

- Current version
 - 29,000 lines of code including comments
 - 68 different programs (scripts)
 - 480 functions in 70 libraries
 - From SLOCCount
 - 19,500 lines of code (no comments)
 - Estimated 4.5 person-years to develop
 - Estimated cost to develop: \$610K
 - Clearly overestimated, but interesting



DocDB Version Numbers

- We use x.y.z version numbers
 - X: Underlying DB format changes (also usually major changes)
 - Y: Major feature changes users are likely to notice
 - Z: Minor feature changes and bug fixes
- Since DB format changes are disruptive, we want to keep these to a minimum



DocDB Version 7

- Current version is 7.2, soon to be 7.3
 - So far, version 7 has
 - Split events from topics (a hack from 1.0)
 - Added a Calendar
 - Allows “Related Documents” to have versions and be in other DocDBs
 - Allows users, meeting organizers, and admins to change which columns are shown when viewing documents, agendas (in 7.3)



Possible V7 Improvements

- There is still room to improve Version 7
 - **Events**
 - Notification when documents added
 - Email reminders of an event
 - **Custom lists**
 - Embed choices in URL
 - Change sort order, re-sort a list



Possible V7 Improvements

- Other possible V7 improvements
 - **XML Insert/Show Document to complement API**
 - Easier insertion of docs by external programs
 - Easy access to DocDB info by ext. programs
 - Would allow external DocDB “Related Docs” to show titles, not just numbers
 - **Web based configuration**
 - Lots of options in DocDB/Project Globals
 - Most can be stored in DB (not security related)



Version 8 Thoughts

- DocDB still doesn't scale up very well for large, disparate groups (more than 300-400 people)
- I think a few changes can help with this
 - This would touch on three major areas in DocDB
 - Topics
 - Events
 - Groups



V8 Topic Thoughts

- Allow infinite hierarchy of topics
 - MajorTopics ➔ Top Level
 - MinorTopics ➔ Second Level
 - Then can have SubSubSub...Topics
 - Would allow documents to have top level topics as topics on document (this has been a request/confusion). Would also be able to set a minimum level which can be applied.



V8 Event Thoughts

- In a large meeting, it can be difficult to find the sessions you want
 - Meetings and/or sessions within meetings can have optional topics
 - Lists of topics will also list events with those topics
 - Also allow certain sessions agendas to always be shown in event agendas (e.g. plenary sessions)



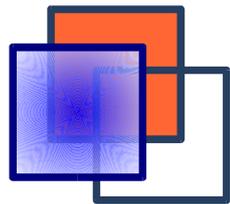
V8 Group Thoughts

- An old idea: Various groups do not see all options
 - e.g. certain groups may not be able to create documents with certain topics
 - or even see that a document HAS that topic
 - or have to take an extra step to add such a topic
 - same with events, document types, etc.
 - Allows moderators to designate “official” docs



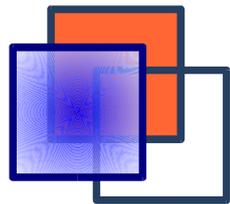
Other Minor Changes

- Requester will be renamed submitter to reduce confusion
- Allow it to be omitted when just one author
- “# of files” on “change page” will be filled in to lessen confusion
- Instructions will explain use of Javascript, popups, and cookies
 - These are used sparingly, only when needed, and DocDB is functional without them



Getting a Change Made

- Since the cancellation of BTeV, I don't do much DocDB work gratis (calendar is one exception)
- So
 - Make the change yourself (contact me first, I do have coding guidelines (which evolve))
 - Convince CD it's a good enough idea to ask for
 - Contract to change it (I may offer “licenses” with programming time included)
 - Make a trivial or obvious request



Getting a Change Made

- In any case...
 - Changes must take DocDB in what I feel (or can be convinced) is the right direction
 - Ideally, everyone would benefit from your enhancement
 - Your change cannot make life more difficult for those who don't benefit from it
 - I'm not interested in incorporating other technology into DocDB (e.g. Plone, InDiCo, etc.) but DocDB could be used as a backend to those systems with an XML (or other) API



How Do I?

- Change options in DocDB
 - Any setting in DocDBGlobals.pm can be changed in ProjectGlobals.pm
 - More useful ones:
 - `$Preferences{Options}{DynamicFullList}`: No cron jobs
 - `$Preferences{Options}{AlwaysRetrieveFile}`: No 2nd password
 - Not default since it might not give same MIME types as Apache
 - `$EnhancedSecurity`: Separate lists of viewers/modifiers
 - Will be default at a later point (DB entries are backwards compat.)
 - `$UseSignoffs`: Allow signoff document control
 - `$ReadOnly`: Password protected, RO version of DocDB



How Do I?

- Search contents of documents?
 - See `scripts/ContentSearch.template`
 - Set `$ContentSearch` to the name of your script
 - Example provided for `htdig`
- Add links to the front page
 - See `ProjectMessages.pm`
 - Examples provided for conference documents, meeting group documents, publications
 - Used by BTeV, Beams for Old -> new mapping



How Do I?

- Add Documents to DocDB not using DocDB
 - See AddDocument in DocumentUtilities.pm
 - Ask on the list for examples of how to call AddDocument
- Several example scripts in scripts and scripts/examples
 - Resecure: Reset PW based security after changing groups
 - ScriptInstall: for managing several instances at once
 - Change security based on keywords



Conclusions

- BTeV's cancellation has slowed development
- But, development does continue and adoption is continuing to grow
- Also, the internal code quality continues to improve (making future developments easier/possible)
- Thank you all for coming, and I welcome your input



Demo of Version 7.3

- Demo
- Other presentations
- Questions