

Unravelling the Pain of Collaboration: LIGO Challenges and Experiences

Scott Koranda for LIGO

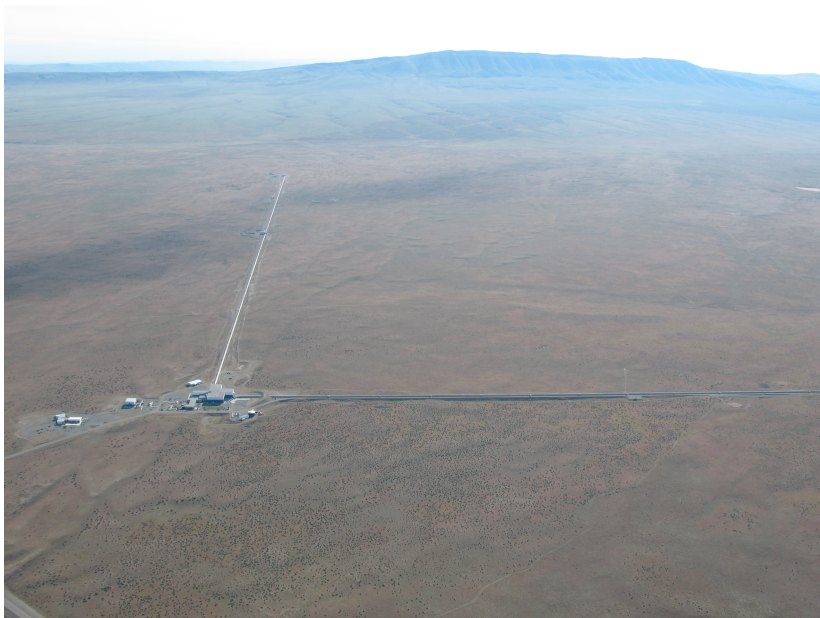
LIGO and University of Wisconsin-Milwaukee

October 5, 2011
LIGO-GXXXXXX-vN



LIGO, the Laser Interferometer Gravitational-wave Observatory, seeks to detect gravitational waves – ripples in the fabric of spacetime. First predicted by Einstein in his theory of general relativity, gravitational waves are produced by exotic events involving black holes, neutron stars and objects perhaps not yet discovered.

LIGO Hanford, WA



LIGO Livingston, LA



LIGO Laboratory =
LIGO Caltech + LIGO MIT +
LIGO Hanford Observatory +
LIGO Livingston Observatory

The LIGO Scientific Collaboration (LSC) is a self-governing collaboration seeking to detect gravitational waves, use them to explore the fundamental physics of gravity, and *develop gravitational wave observations as a tool of astronomical discovery*. The LIGO Scientific Collaboration was founded in 1997 and currently has more than 800 members from 70 institutions worldwide.

LIGO LIGO Scientific Collaboration LSC



Started in Summer 2007

Knit together existing technologies and tools

Goals:

- Single identity for each LIGO person
- Single source of membership info
- Single credential for each LIGO person
- SSO across web, grid, command-line

Found we had two building blocks:

- ① The nascent “LIGO Roster” project
 - PHP + Apache + MySQL
- ② Kerberos principal for each LIGO member
 - unused at the time
 - `scott.koranda@LIGO.ORG`
 - users call it their “at LIGO.ORG login”
 - also known as their “albert.einstein” login

Good: Branding “albert.einstein” and “@LIGO.ORG”

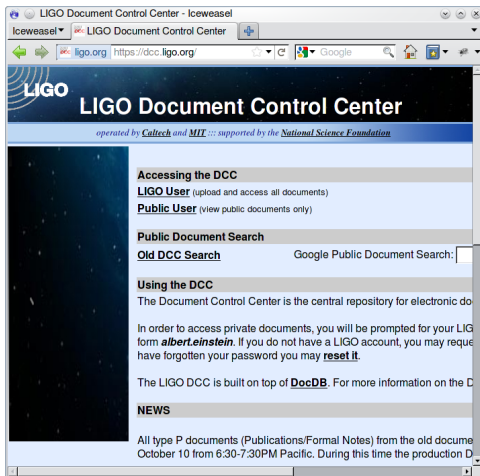
- Users understand what's being asked for
- Thought leaders buy in faster and help promote
- Managers buy in faster and help promote
- Admins focus easier on delivery
- Users push admins more to deliver

Why can't I access this site with my LIGO.ORG?

Good: Find the killer app for CO

LIGO killer app is DCC:
document control center

- All scientific papers managed through DCC
- Can't check name on paper if can't access DCC!
- We saw quick adoption of new LIGO identity
- Later the wikis also become killer app



(Why not data access as killer app? We could not perturb scientific progress...)

Selecting solid tools was the easy part...

Kerberos single identity and credential

Grouper single source of membership info

OpenLDAP solid distributed replication

Shibboleth web SSO with eye towards federation

Sympa sophisticated email list management

COmanage customizable CO management (coming soon!)

Good: Kerberos principal as identity

- Solid, well weathered protocol
- Exchange easily for other tokens
- Password strength checking
- Cross web, grid, command-line boundaries
- Distributed service just works

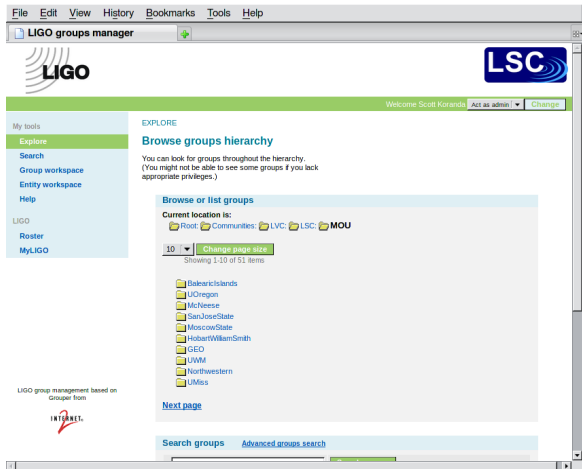
Compare to an OpenID approach...

Bad: Kerberos principal as identity

- Slow evolution
- MS Active Directory is not "just Kerberos"
- SPNEGO is great until it isn't
- (too many ISPs block port 88)

Good: Grouper for membership info

- Inheritance simple and scalable
- No assumption of structure
- Reflection into LDAP
- Web services interface
- Permissions baked in

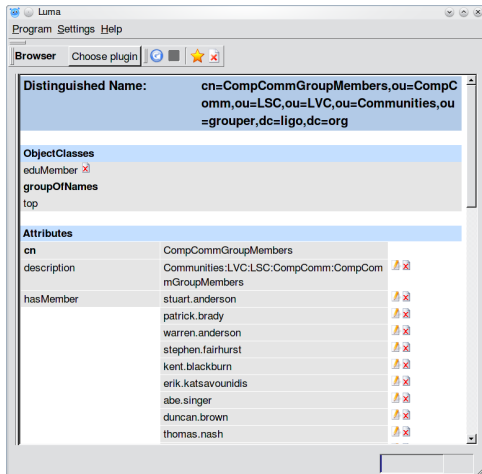


Bad: Grouper for membership info

- Non trivial deployment exercise
- No namespace structure—get out ahead of it!
- Default UI not suitable for non-experts
- “Lite” UI better but still not CO-specific

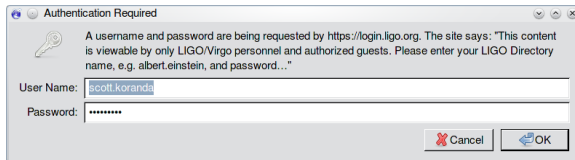
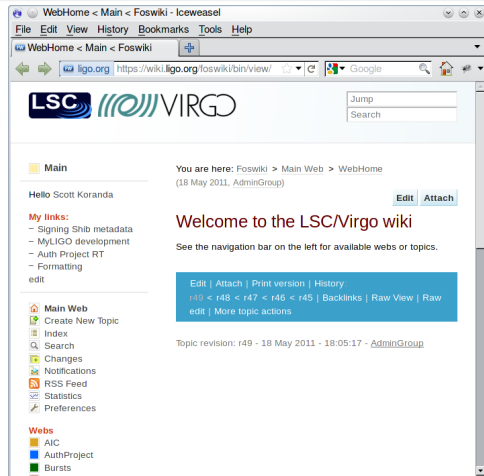
Good: OpenLDAP for solid distributed replication

- Use latest 2.4.x stable version!
- Makes consuming membership and attributes easy
- Just works



Good: Shibboleth for web SSO

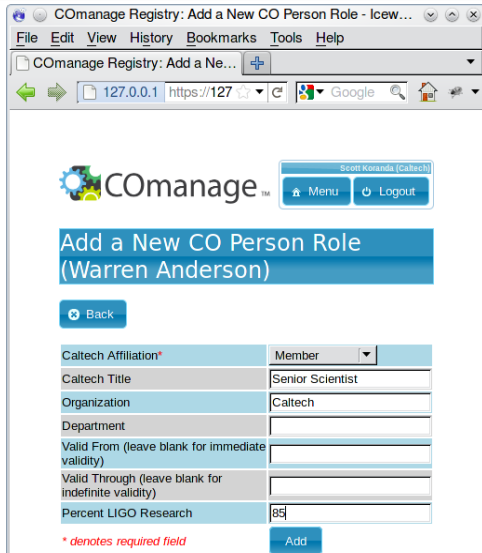
- Scalability—unlikely LIGO will push scaling
- Feature rich—found little we cannot accomplish
- Extensibility—customization for edge cases, ECP
- Federation—evolution path for LIGO is clear



- Science COs not in the “sweet spot”, campus federation drives agenda
- Example: IdP high availability across WAN
- SAML2 learning curve significant for architects, integrators
- (not so bad for admins)

Good: CManage

- Easy collaboration management
- Customizeable enrollment flows
- Extensible CO specific attributes
- Batteries included! (Registry)
- LIGO contributing directly



The screenshot shows a web browser window titled "CManage Registry: Add a New CO Person Role - Icew...". The browser's address bar shows "https://127.0.0.1:127...". The page features the CManage logo and a user profile for "Scott Koranda (Caltech)" with "Menu" and "Logout" buttons. The main heading is "Add a New CO Person Role (Warren Anderson)". Below this is a "Back" button and a form with the following fields:

Caltech Affiliation*	Member
Caltech Title	Senior Scientist
Organization	Caltech
Department	
Valid From (leave blank for immediate validity)	
Valid Through (leave blank for indefinite validity)	
Percent LIGO Research	85

* denotes required field

Add

It's not a released product yet!

(LIGO planning to leverage a March 2012 release)

...hard part is in the details for our CO

- building the UIs for basic identity management
- policy debate blackholes
- rediscovering lessons learned by others
- domestication ain't done 'till it's done
- no corner case shall go unexplored
- provisioning and de-provisioning
- highly distributed community
- federating with smaller COs

Ugly: Building CO-Specific UIs

- No dedicated FTE for UI design and implementation
- Small pool of talent and resources to draw on
(NSF does not fund us to develop identity management UIs)
- Ever evolving requirements and use cases and corner cases
- High user expectations (if Google can...)
- Weakness and brittleness in identity management UIs propagates and poisons rest of infrastructure

Ugly: Policy Debate Blackholes

- Does FERPA apply to LIGO? If so, how?
- What is required for managing demographics?
- ADA, DDA in the UK, German privacy laws?
- Opt-in, opt-out, information backup & archiving, logging,...

We're just trying to detect gravitational waves!

LIGO doesn't have expertise or experience to easily address these types of issues when they come up.

Ugly: Rediscovering Lessons Learned By Others

- People get married and change name
- People change gender and change name
- People leave and come back
- People leave and come back with different name
- People leave and join a federating CO

Ugly: Domestication Ain't Done 'till It's Done

LIGO cannot take a greenfield approach

We domesticate the tools already being used:

- TWiki/Foswiki: must hack the Perl code
- MoinMoin: must hack the Python code
- Dokuwiki: must hack the PHP code
- eLOG: must hack the PHP code

Where is the roadmap for domesticating “simple” tools like wikis?

So busy domesticating legacy apps little time to explore Foodle, Google, other solutions

Ugly: No corner case goes unexplored

Requirements for the MoinMoin wiki at UWM:

Most pages should require authentication to view, and LIGO users should use @LIGO.ORG credentials to authenticate, but UWM users should use their ePantherID, and still others should be able to use once-off passwords we give them. Some pages should be viewable by the public, and fine-grained ACLs should work for everything.

Doable, but don't underestimate the FTE cost!

Ugly: Provisioning and De-provisioning

Primary concern is LIGO Data Grid computing cluster accounts

10 different sites, each managed independently

- Is an automated approach even technically possible?
 - uid/gid mapping via NIS+, LDAP, plain text files
 - different storage models at each site
 - different file systems at different sites
- Would local admins allow automated account creation?
- How should a 10 TB home directory be de-provisioned?
- When should de-provisioning happen?
 - When Important Person signs off?
 - At a specific date and time?
 - Just how many corner cases can we discover?
- How can we be sure it happens at the appointed time?

Ugly: Highly Distributed Community

- Pls only way identities and membership vetted
(some Pls can't be bothered to keep roster current)
- services distributed and ephemeral—coordination difficult
- some providers lack expertise, have no campus IT backstop
- still identifying our sticks and carrots...

Ugly: Federating With Smaller COs

Real (and in some cases urgent) science drivers

Need collaboration spaces with

- astronomers
- astrophysicists
- numerical relativists

Requirement to leverage existing @LIGO.ORG identities

Smaller COs have no (managed) identities to offer!

Until everyone federated try to offer @LIGOGUEST.ORG identities

Plumbing isn't the problem...it's all the UI and process work