Identity Management for the LIGO Project

Scott Koranda for LIGO

LIGO and University of Wisconsin-Milwaukee

April 28, 2010
LIGO-G1000471-v3
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.
Who we are...

(’cause it’s complicated and puts demands on our tools)
LIGO Hanford, WA
LIGO Livingston, LA
LIGO Caltech
LIGO Laboratory

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 nearly 800 members from 70 institutions worldwide.
LIGO Scientific Collaboration

[Logos of various universities and institutions]
The LSC is an open collaboration. Anyone who is interested in contributing to the mission of LIGO may apply to join the LSC. Groups joining the LSC are welcome to participate in all LSC activities. As members of the LSC, groups have access to the LIGO data for scientific purposes, can participate in collaboration meetings and working group meetings, can represent the LSC at external scientific meetings and to the public, and have representation on the LSC Council, the governing body of the collaboration.
LIGO Scientific Collaboration

...groups prepare and sign Memoranda of Understanding (MOU) with the LIGO Lab and the LSC outlining the group’s role in the LSC as well as specific research plans for the coming year...
Some, *but not all*, members of the LIGO Lab are members of the LSC.
“Groups” join the LSC
(some groups more structured than others)
GEO600 interferometer, Hannover, Germany
GEO600 Members

- Gravitational Physics Group, University of Birmingham
- Gravitational Physics Group, Cardiff University
- Institute for Gravitational Research, University of Glasgow
- Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Potsdam
- Max-Planck-Institut für Gravitationsphysik, (Albert-Einstein-Institut), Hannover
All members of the GEO600 project are members of the LSC
On the other end of the spectrum…

CalState Fullerton LSC Group

(Josh)
LSC or LIGO?

Internally and almost always when presenting our external face we simply call ourselves “LIGO”
Broader GW Community

GW community is larger than LIGO...
Virgo interferometer, Cascina, Italy
Virgo and the LSC

Virgo members are *not* members of the LSC
Virgo and the LSC

Virgo and LIGO...

▶ share access to data
▶ share access to computing resources

Joint body is “LIGO/Virgo Community” or LVC
Why is membership important?

- access to data
- names on papers

Two items scientists care about *intensely*
Today...

- 810 current and active members
- Single authoritative roster of members
- Single LIGO identity for each member
How we got here

It wasn’t always this way...
The mess we made on the Grid

LIGO Data Grid (LDG)

- 15000+ cores
- 10 sites
- Many flavors of data and metadata services
- > 200 users
The mess we made on the Grid

- LDG emerged in 2001
- Sought single sign-on and promise of Grid utopia
- Most Grid tools require PKI and GSI
The mess we made on the Grid

- User must request, retrieve, manage X.509 cert
  - Not all web browsers do PKI well
  - Grid tools require PEM but web browsers write PKCS12
  - “17, but steps 6) have 9) have 12 or 13 subitems each”
  - Turns out Ph.D. physicists on average cannot do this
  - Command line tools don’t help much
The mess we made on the Grid

- No roster of who is/is not member of LIGO
- Each cert request must be vetted
  - Requires “secure communication” with each group PI
  - Getting attention of PIs can be difficult
  - SMIME email difficult for most PIs
  - Loop not closed when people leave group
The mess we made on the Grid

After X.509 cert issued user must be authorized
  ▶ Cumbersome
    ▶ Each user added by hand to ACL file(s) at each site
    ▶ Only grid-specific solutions available for managing ACLs
  ▶ Not uncommon for new member to wait weeks for credentials and access to LDG resources
The mess we made on the Grid

Managing access to LDG was one of the first hints we needed better identity management...

...we didn’t take the hint...
The mess we made on the Web

- Early use case: eLogs at the sites
  - Web based electronic notebooks
  - Email “the” admin for access (hopefully he knows you)
  - Unique accounts, but...
  - All accounts use the same password
  - Loop not closed when people leave collaboration
The mess we made on the Web

- Multiple sites deploying web tools
  - GNATS, Bugzilla, Redmine, Trac, Gitorious?
  - Moin, Twiki/Foswiki, Docuwiki, MediaWiki,...
  - Each requiring new login/password for user
The mess we made on the Web

Users frustrated

First response is “well known login/password”

- shared login and password collaboration wide
- used for protecting “low risk” information
- who monitors what is low risk?
- found login/password in the wild
The mess we made on the web

As the number of web tools and services grew we knew we had a problem...

...but we were in production, busy doing science, and didn’t take the hint...
The mess we made of Email

mailman is *not* a collaborative tool

- Each list admin needs to add people individually
- Archives require yet another login/password
- People change institutions and addresses
- Members leave collaboration but stay on the lists
The mess we made on the command line

Version control repositories

- CVS, SVN, git
- Distributed across multiple sites
- Each requiring yet another login/password
- People leave collaboration but still have access

Same issues for other command line tools
Managing access for hundreds of people to multiple code repositories was a nightmare...we knew we had a problem...

..but we were in production, busy doing science, and couldn’t take the hint...
We had a mess

- No single event precipitated new approach
- It really came down to two things:
  1. Sustained whining from frustrated users
  2. Chatting with Ken Klingenstein over drinks
LIGO Identity Management Project

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
LIGO Identity Management Project

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
   ▶ roster drives creation of principal for each member
   ▶ roster pushes principal and details into LDAP
Single authoritative source of membership

Decided to leverage Grouper from I2

- Flexible enough to reflect community structure
- Ready-to-use web front-end
- SOAP and RESTful WS APIs
- Privilege support
- Reflect into LDAP
EXPLORE

Browse groups hierarchy
You can look for groups throughout the hierarchy. (You might not be able to see some groups if you lack appropriate privileges.)

Browse or list groups
Current location is:

- Root
- Communities: LVC: LSC: MOU

10 ▼ Change page size
Showing 1-10 of 51 items

- BalearicIslands
- UOreogn
- McNeese
- SanJoseState
- MoscowState
- HobartWilliamSmith
- GEO
- UWM
- Northwestern
- UMiss

Next page

Search groups Advanced groups search
My memberships

To find groups in which you are a member, you can:
- Browse the groups hierarchy
- List your groups
- Search for groups by name

Browse or list groups

Showing 1-7 of 7 items

- Communities:LVC:LSC:CompComm:AuthProject:AuthProjectGroupMembers
- Communities:LVC:LSC:CompComm:CompCommGroupMembers
- Communities:LVC:LSC:LSCGroupMembers
- Communities:LVC:LSC:MOU:UWM:UWMGroupManagers
- Communities:LVC:LSC:MOU:UWM:UWMGroupMembers
- Communities:LVC:LVCGroupMembers
- Grouper Adminstration:SysAdmin Group

Search groups

Search groups

Display results by

Path Name ID Path
Leveraging Grouper

- MOU membership out of “roster” into Grouper
- Use roster PHP to drive Grouper SOAP WS
- Subject/member info pulled from LDAP
- Memberships pushed into LDAP (bushy)
[root@oregano ~]# ldapsearch -LLL -b "ou=people,dc=ligo,dc=org"
-H ldap://ldasdata4.ligo.caltech.edu -x '(cn=Scott Koranda)'
isMemberOf
dn: employeeNumber=882,ou=people,dc=ligo,dc=org
isMemberOf: Communities:LVC:LSC:MOU:UWM:UWMGroupMembers
isMemberOf: Communities:LVC:LVCGroupMembers
isMemberOf: Communities:LVC:LSC:LSCGroupMembers
isMemberOf: Communities:LVC:LSC:CompComm:CompCommGroupMembers
isMemberOf: Communities:LVC:LSC:MOU:UWM:UWMGroupManagers
LIGO Roster

- Students, post-docs, can apply for membership
- Managers approve & add/remove members
  - Access control derived from Grouper privileges
- Members manage password for LIGO identity (Kerberos principal)
LSC Member Management

LSC Group:

LSC - UW Milwaukee

Act on Pending Membership Requests
There are currently no pending membership requests for this group.

Actions:

- Act on Pending Membership Requests
- Manage Members
- Manage Council Delegates

LIGO Directory Services and Authentication & Authorization Services Infrastructure are supported by the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.
Single identity and authoritative membership is key

LIGO Roster, Grouper, and Kerberos a powerful combination

- Kerb principal enables single identity
- Roster enables management of those identities
- Grouper enables management of memberships

With this foundation we could tackle web, grid, and command line spaces...
Single sign-on for LIGO web space

Deploy I2 Shibboleth System

- Single sign-on across LIGO web tools/pages
- LIGO Identity Provider (IdP)
  - Authenticate via REMOTE_USER and mod_auth_kerb
  - Attributes pulled from LDAP master server
  - Focus mainly on IsMemberOf (via Grouper)
- Look to federate in future
  - InCommon for many U.S. institutions
  - European federations (UK, DFN-AAI)
  - Virgo?
A username and password are being requested by https://login.ligo.org. The site says: “This content is viewable by only LIGO/Virgo personnel and authorized guests. Please enter your LIGO Directory name, e.g. albert.einstein, and password…”

User Name: 

Password: 

Cancel  OK
Welcome to the LIGO Data Grid wiki which compliments the LIGO Data Grid and DASWG websites.

Organizational

- Telecons and Minutes
- LIGO Data Grid Team
- Version Control Systems Advisory Committee (2008)
- OS Selection Committee (2007)

Projects

- S6 Online Working Group
- Advanced LIGO Data and Computing
- GPU Development in LSC/Virgo
- CondorC for LIGO Data Grid
Center for Gravitation and Cosmology Wiki

Contents

1. Administrative
2. Projects/Events
3. MOU Reports and Attachments
4. Computing
5. Proposals
6. General HowTo's URL
7. UWMLSC Document Center

Administrative

- Information for new group members
- Useful links and resources for LSC members
- Information for travellers
- Information for visitors

Projects/Events

- Schedule of Departure/Return times for LIGO/VIRGO meeting 2010
- projects/GWDAW15

MOU Reports and Attachments
- Only deploying Shibboleth 2.1
- Relying heavily on SAML2/Artifact
- Helps with a specific use case
  - Wiki include IMG from multiple servers
  - Browsers won’t “do right thing” when fetching IMG
  - SAML2/Artifact requires no JS
L1:OMC-QPD1_Y_OUT_DAQ and L1:OMC-QPD2_Y_OUT_DAQ have some noise at higher frequencies:

LSC-DARM_ERR  OMC-QPD1_Y_OUT_DAQ  OMC-QPD2_Y_OUT_DAQ

---

H1: There is a H1:DMT-PRE_LOCKLOSS_1800_SEC flag.

Looks like some H1:OMC-QPD1 channels had some noise at higher frequencies:

LSC-DARM_ERR  H1:OMC-QPD1_Y_OUT_DAQ
MyProxy and GridShib CA integrate LIGO Data Grid

- MyProxy exchanges Kerb ticket for X.509 cert
- GridShib CA exchanges SAML2 for X.509 cert
- User “sees” @LIGO.ORG cred required for both
- X.509 certs are “short-lived”
- Can also be converted to RFC 3820 proxy cert
LIGO Certificate Authorities

MyProxy and GridShib expose LIGO CA

- SLCS = short lived credential service
- The Americas Grid Policy Management Authority (TAGPMA)
- TAGPMA provides SLCS profile
- Plan to accreditate LIGO SLCSs
LIGO Data Grid Authorization

Grid authorization driven by Grouper and LDAP

- ACL files derived from IsMemberOf
- Simple LDAP query with local caching
- X.509 DN and login pulled from LDAP

```
# isMemberOf Communities:LVC:LSC:LSCGroupMembers
"/DC=org/DC=ligo/O=LIGO/OU=People/CN=scott.koranda@LIGO.ORG" skoranda
"/DC=org/DC=ligo/O=LIGO/OU=People/CN=sergey.klimenko@LIGO.ORG" klimenko
"/DC=org/DC=ligo/O=LIGO/OU=People/CN=sukanta.bose@LIGO.ORG" bose
"/DC=org/DC=ligo/O=LIGO/OU=People/CN=szabolcs.marka@LIGO.ORG" smarka
"/DC=org/DC=ligo/O=LIGO/OU=People/CN=teviet.creighton@LIGO.ORG" teviet
```
Integrating the command line

CVS, SVN, git tunnel through SSH
  ▶ Most Linux OpenSSH sshd GSS-API + Kerberos
  ▶ Grid-enabled OpenSSH also deployed
  ▶ NCSA “mechglue” enables Kerb + GSI
  ▶ PAM also work with Kerberos

This pattern same for other command line tools

(note that curl works well with SAML2/Artificat)
Integrating email

- LDAP queries define lists
- Fairly complex queries possible
- `mailAlternateAddress` LDAP attribute enables posts from multiple accounts
- Lists can accept posts from any person in collaboration
- Web access to list management pages and archives via Shib
Putting it all together

Within 15 minutes of joining LIGO Albert Einstein using his albert.einstein@LIGO.ORG credential can...

1. Access LIGO wikis to find HOWTOs
2. Download and install client tools
3. Login to cluster
4. Checkout code from git repository
5. Email analysis discussion list for help
6. Build code, submit analysis jobs

From 0 to science with one @LIGO.ORG credential
Cleaning up is easy

When Albert Einstein leaves the LIGO collaboration...

1. albert.einstein@LIGO.ORG Kerberos principal disabled
2. Removed from Grouper/LDAP groups
3. No login to Shib IdP, no web access
4. No MyProxy, GridShib, no grid access
5. No access to code repositories
6. No email lists
So everything is perfect?
The problem of single sign-on

- Compromised @LIGO.ORG credential yields much access
  - Users don’t do great job protecting credential
- Do we want to enable access all way to instruments?
  - Probably require N-factor auth closer to instrument
- Federation: should the same ID faculty use for managing grants and grading be used to access LIGO resources?
  - We will have to get better at levels of assurance
Yes, we would like single sign-off

It’s hard

Nuff’ said
Too many auth popup boxes

Used `mod_auth_kerb` as bridge until Shib deployed

- Wiki pages with 100s of images linked
- Threaded browsers do multiple GET of IMG
- User sees many, many auth popups
- Some browsers just crash
- One interim fix is a special "authentication page" with 1 image from each server
Firewalls and blocked ports

Tried SPNEGO to help users with auth across servers

Bad, bad interaction for Mac OS X users

- Mac keychain caches and manages Kerb ticket
- Hotels, campus guest wireless, many others block port 88 for kinit
- Firefox browser/key chain will not give up
- No web access for users at critical times

Investigating moving Kerberos KDC to listen on port 443
Remaining challenges and needs
Grouper needs

- Support for point-in-time auditing
  - Authorship determined by membership dates
  - “Who was member of LSC on July 1, 2009?”

- Support for multi-value attributes on subjects
  - Addresses, phone numbers, ...
  - Reflect into LDAP
  - Ideally we can dump our own MySQL tables
Shibboleth needs

- **IdP “clustering” across WAN**
  - Less about performance more about robustness
  - IdP login must still function when observatory sites lose network connectivity to outside

- **More turnkey SP integration with web tools**
  - Moin, Twiki/Foswiki, Dokuwiki
  - Only find hints and half-baked solutions
  - Penetration into those developer communities to help embrace?
Distinction between web and grid is fading

- Scientists just want to use tools
- Don’t care if “web” or “grid”
- Typical use case:
  - Submit large workflow to grid
  - Jobs run for week analyzing data
  - Workflow generates 1000’s of summary images
  - Need to POST summary into analysis wiki
- Seamless cred management across grid, web, cloud
- Delegation is important
  - Workflow management systems need to cache and refresh credentials during lifetime of workflow
  - LIGO working closely with UW Condor team
- Need I2 and grid communities to build together
I2 technologies enabling science

Grouper and Shibboleth are not just for campus IT anymore!

Expect I2 software to impact more science VOs going forward