LIGO Identity Management: How a Scientific Software Security Innovation Institute (S3I2) Might Have Helped

Scott Koranda for LIGO

LIGO and University of Wisconsin-Milwaukee

October 26, 2011
LIGO-G1101185-v2

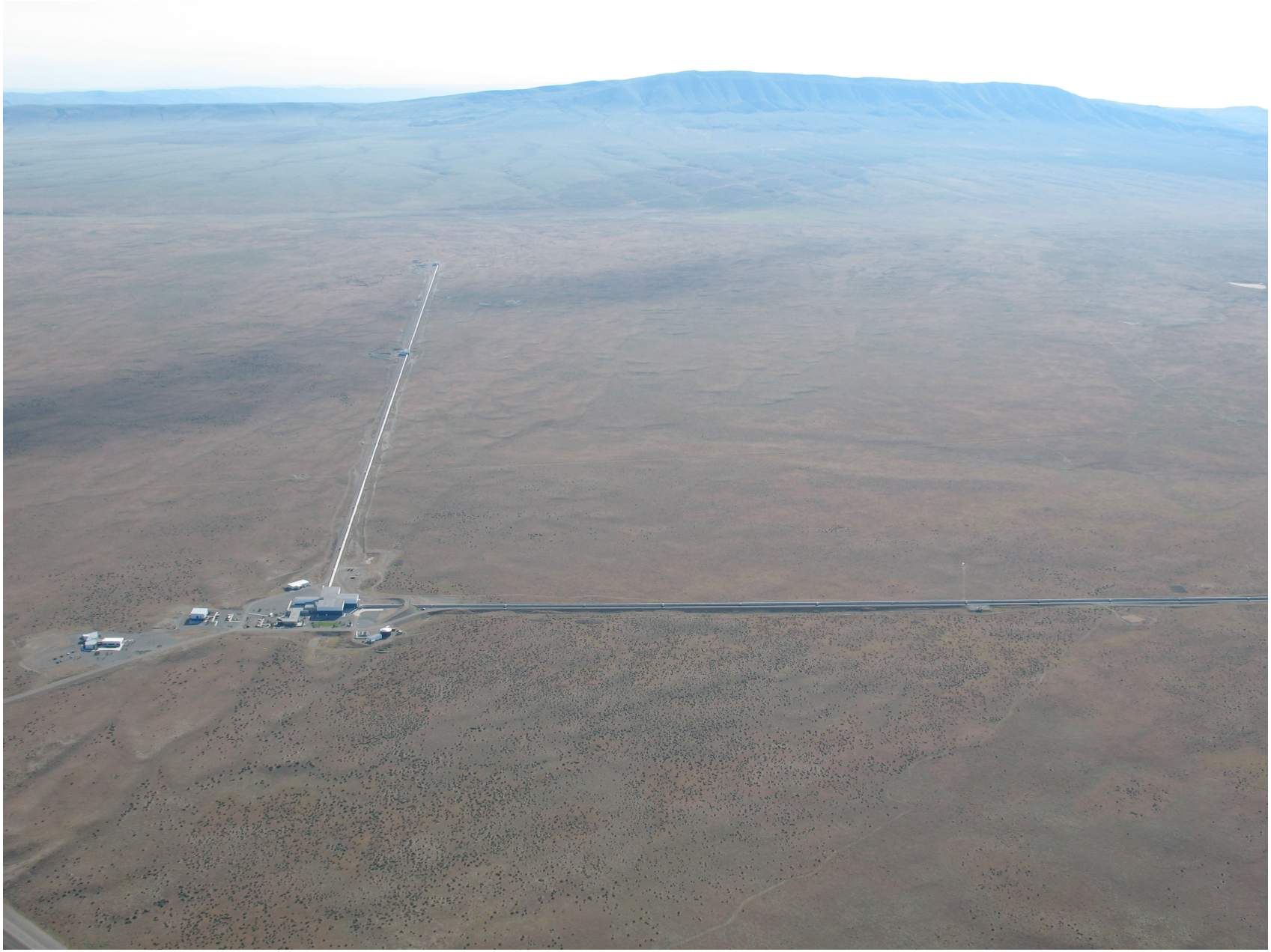
LIGO Science Mission

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

LIGO Scientific Collaboration

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 over 800 members from more than 70 institutions worldwide.


LIGO
LIGO Scientific Collaboration
LSC
University of Maryland
Michigan
University of Strathclyde
ACIGA
Montana State University
PAH
Loyola University New Orleans
University of Washington
Syracuse University
Tsinghua University
Trinity University
San José State University
University of Wisconsin UWMilwaukee
ANU
The Australian National University
LSU
UTB TSC
The University of Western Australia
University of Glasgow
NASA
Goddard Space Flight Center
Andrews University
California Institute of Technology
Northwestern University
R·I·T
The University of Adelaide Australia
Moscow State University
Embry-Riddle Aeronautical University
University of Southampton
Penn State
Geo 600
Cardiff University
California State University Fullerton
Washington State University
UF University of Florida
Universitat de les Illes Balears
Charles Sturt University
University of Rochester
NAOJ
Southern University
William Smith College
University of Minnesota
Universität Hannover
Science & Technology Facilities Council
Rutherford Appleton Laboratory

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

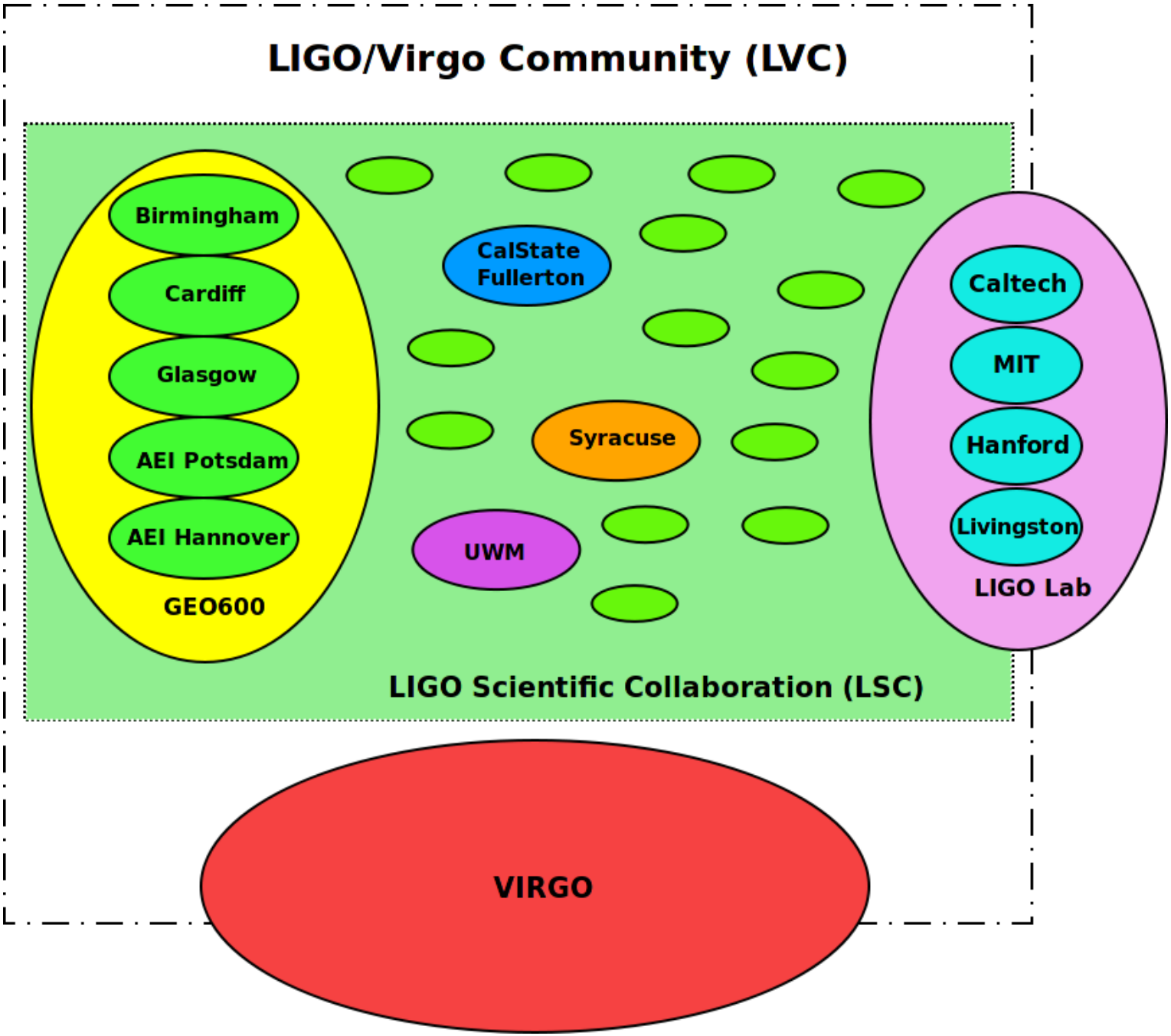
Virgo members are *not* members of the LSC

Virgo and the LSC

Virgo and LIGO...

- share access to data
- share access to computing resources

Joint body is “LIGO/Virgo Community” or LVC


LIGO/Virgo Community (LVC)
Birmingham
Cardiff
Glasgow
AEI Potsdam
AEI Hannover
GEO600
CalState Fullerton
Syracuse
UWM
Caltech
MIT
Hanford
Livingston
LIGO Lab
LIGO Scientific Collaboration (LSC)
VIRGO

Why is membership important?

- access to data
- names on papers

Two items scientists care about *intensely*

LSC Today

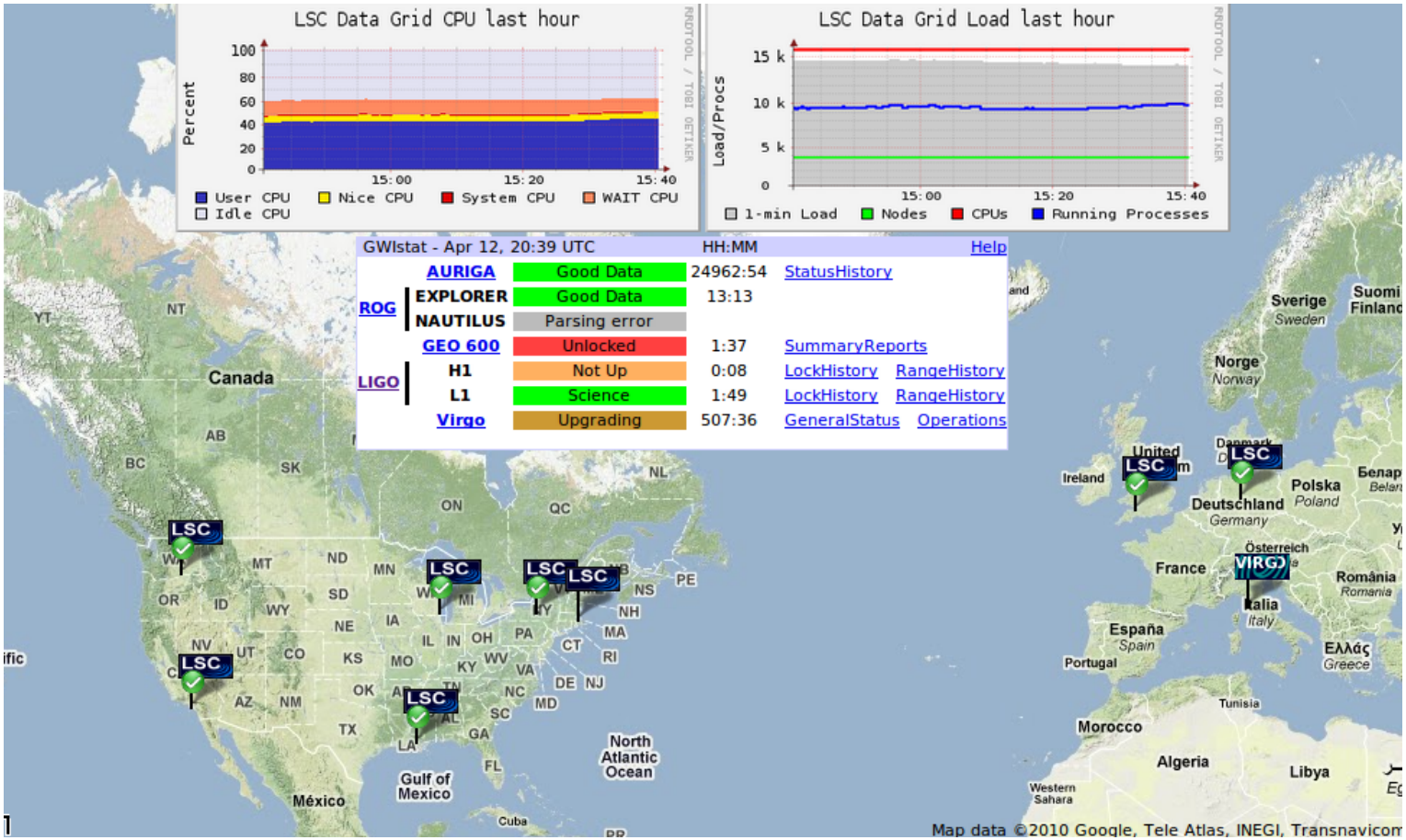
Today...

- 868 current and active members
- Single authoritative roster of members
- Single LIGO identity for each member

It wasn't always this way...

The mess we made on the Grid

LIGO Data Grid (LDG)



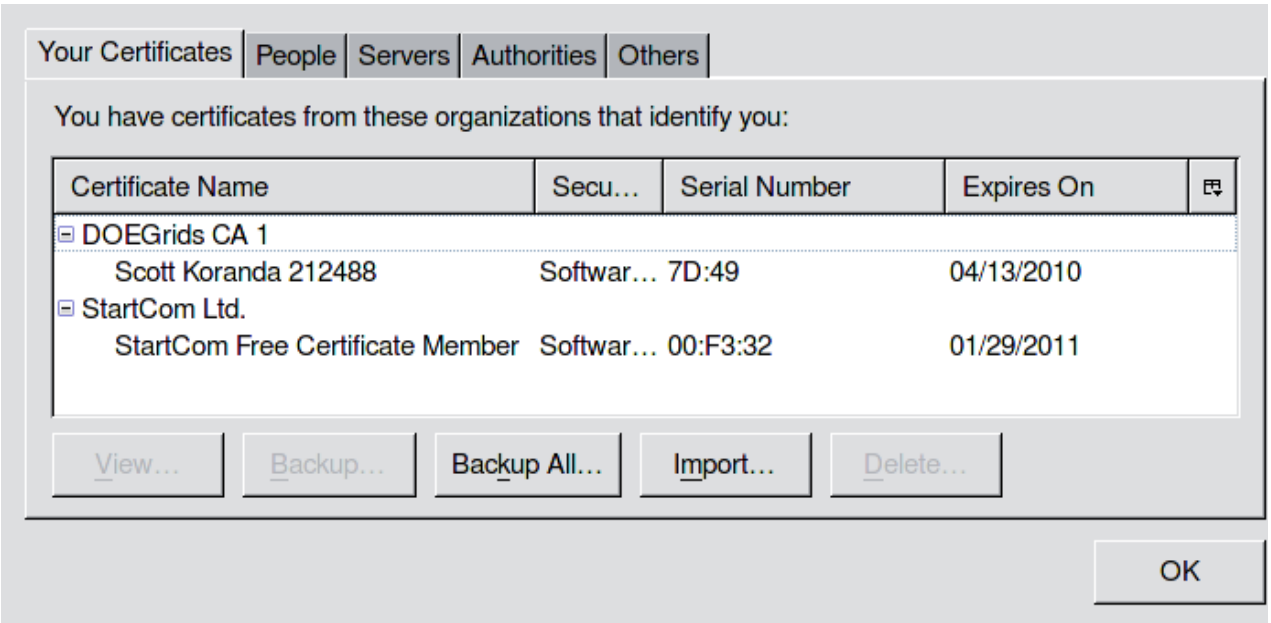
- 15000+ cores
- 10 sites
- Many flavors of data and metadata services
- $>$ 200 users

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
- After cert issued user must be authorized
- Only grid-specific solutions available for managing ACLs

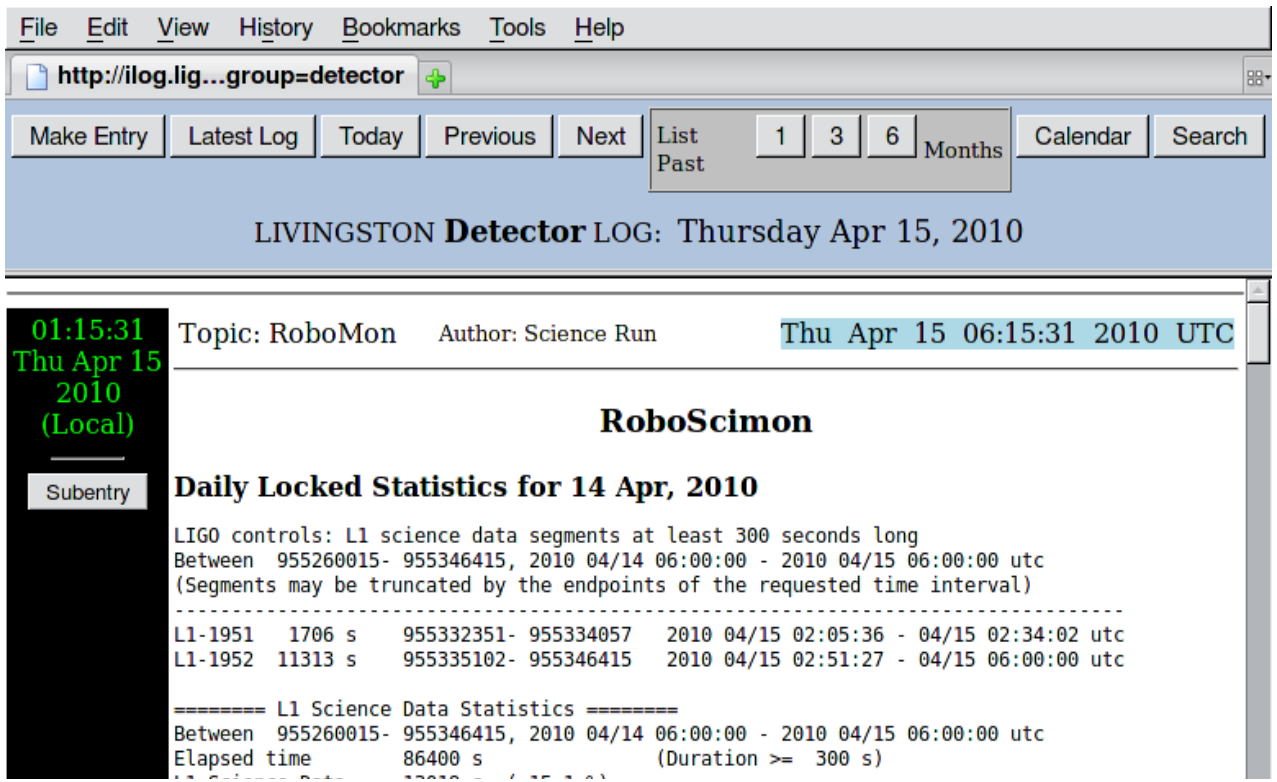
Not uncommon for new member to wait weeks for credentials and access to LDG resources

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

The mess we made on the command line

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

- Late in 2007 and we stopped scaling
- Collaboration business at risk
- No single event precipitated new approach
- It really came down to two things:
  1. Sustained whining from frustrated users
  2. Chatting with Ken Klingenstein (I2) over drinks

Had a NSF S3I2 existed...

We would have asked:

- What is possible for science VOs?
- What would success look like?
- What should our goals be?
- What are the axes of the problem?
- What is the vocabulary for the problem space?
- Who are the players in this space?
- How much will it cost?
- How do we get started?

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

Had a NSF S3I2 existed...

We would have asked:

- Should we build on Kerberos?
- What operational details should we know about Kerberos?
- What password policies should we adopt immediately?
- How do we structure our LDAP?
- Is this PHP + Apache + MySQL approach a good one?

Had a NSF S3I2 existed...

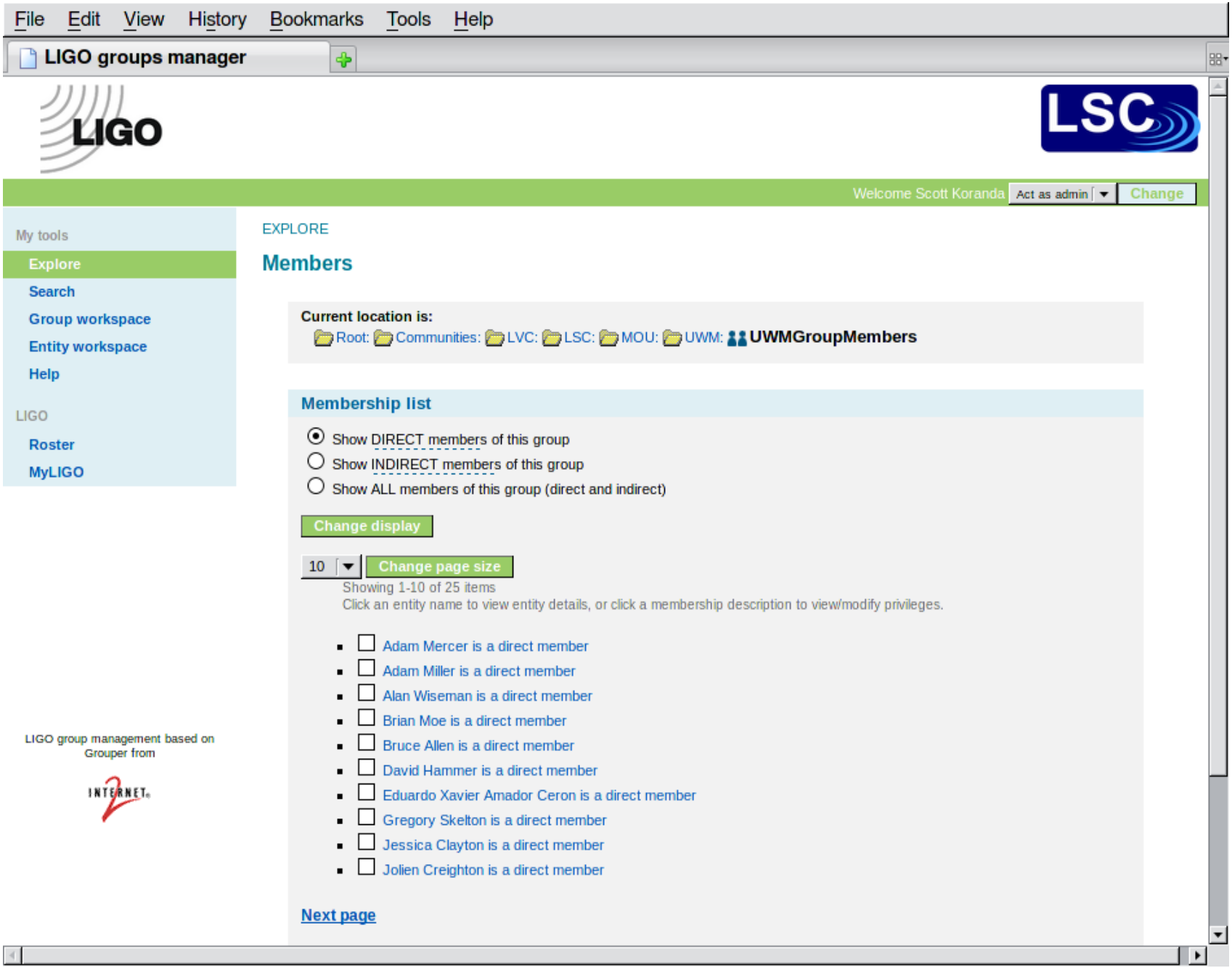
We should have been told:

- Kerberos is good choice for authentication
- Design to separate authentication and authorization
- Do not plan on Kerberos for authorization
- "Here is a solid KDC operations document for science VOs"
- "Here is a best practices KDC policy for science VOs"
- "Here is a best practices LDAP document for science VOs"
- "You need to build a proper registry: the first thing to do is figure out who is in your collaboration, how they enroll (onboard), how they leave (offboard), how identity is managed at a basic level."

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


File Edit View History Bookmarks Tools Help
LIGO groups manager
LIGO
LSC
Welcome Scott Koranda
Act as admin
Change
My tools
Explore
Search
Group workspace
Entity workspace
Help
LIGO
Roster
MyLIGO
LIGO group management based on Grouper from
INTERNET2
EXPLORE
Members
Current location is:
Root: Communities: LVC: LSC: MOU: UWM: UWMGroupMembers
Membership list
Show DIRECT members of this group
Show INDIRECT members of this group
Show ALL members of this group (direct and indirect)
Change display
10
Change page size
Showing 1-10 of 25 items
Click an entity name to view entity details, or click a membership description to view/modify privileges.
Adam Mercer is a direct member
Adam Miller is a direct member
Alan Wiseman is a direct member
Brian Moe is a direct member
Bruce Allen is a direct member
David Hammer is a direct member
Eduardo Xavier Amador Ceron is a direct member
Gregory Skelton is a direct member
Jessica Clayton is a direct member
Jolien Creighton is a direct member
Next page

```
[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
```

Had a NSF S3I2 existed...

We would have asked:

- Should we build on Grouper?
- What is the project arc for Grouper?
- What is the group management ecosystem?
- What is the privilege management ecosystem?
- Namespace?

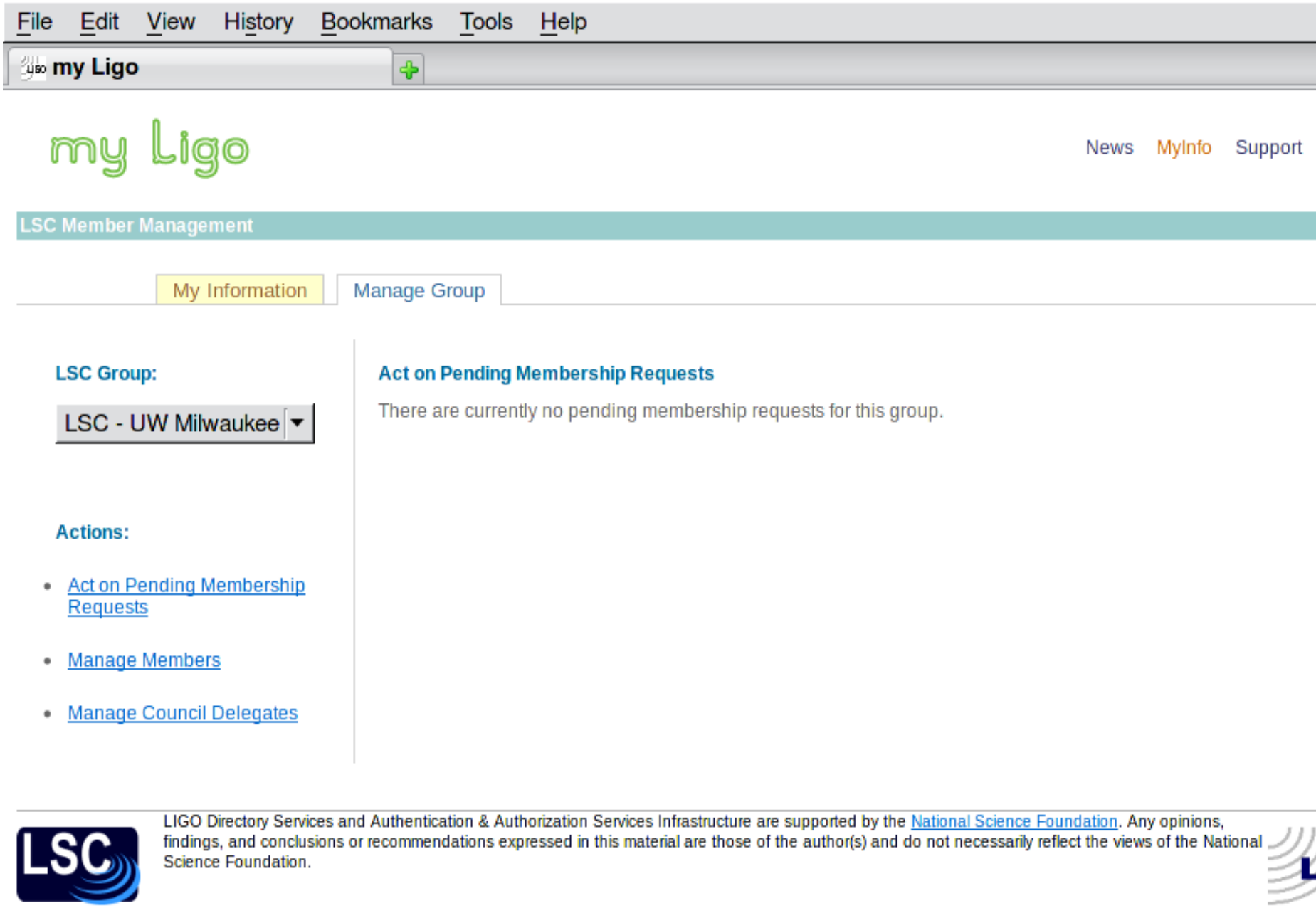
Had a NSF S3I2 existed...

We should have been told:

- Grouper has a solid start but needs 4 years to mature
- Grouper will scale to meet your needs
- Grouper roadmap includes RBAC and privilege management
- "Here is where Grouper fits into the ecosystem"
- "The other tools in this space include..."
- "Here is a group namespace best practices document"

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)


File Edit View History Bookmarks Tools Help
my Ligo
my Ligo
News MyInfo Support
LSC Member Management
My Information
Manage Group
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
LSC
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.

Had a NSF S3I2 existed...

We would have asked:

- Is this MyLIGO approach going to work?

Had a NSF S3I2 existed...

We should have been told:

- "You need to build a proper registry."
- "You need to hire people with these skills:..."
- "The technologies and framework you use is less important than thinking through and documenting clearly how people onboard/offboard and the business processes of your collaboration."

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?

Had a NSF S3I2 existed...

We would have asked:

- Should we build on SAML2 and Shibboleth?
- OpenID? BrowserID? Other?
- Oauth? Oauth2?
- Is federation important? Will it work?
- What role will InCommon play?

Had a NSF S3I2 existed...

We should have been told:

- SAML2 owns Higher Ed
- Internet2 is a major player
- Shibboleth is solid but requires a significant investment
- "Social to SAML" gateways can help you hedge
- InCommon delivers less than you think
- InCommon delivers more than you think
- Federation is important but still in flux
- International federation is still the wild west

MyProxy and GridShib CA integrate LIGO Data Grid

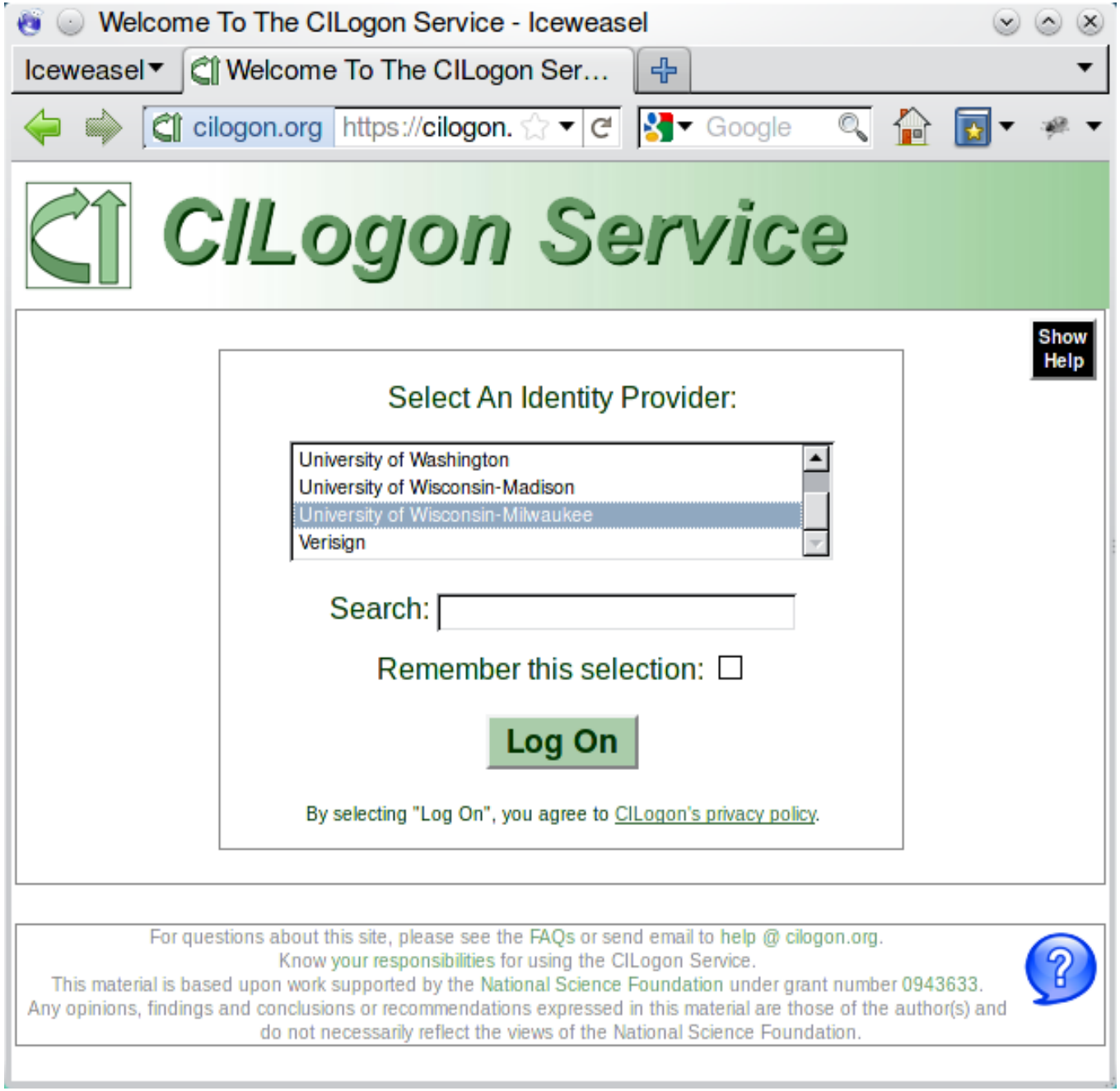
- MyProxy exchanges Kerb ticket for X.509 cert
- GridShib CA exchanges SAML2 for X.509 cert
- User “sees” @LIGO.ORG cred required for both
- X.509 certs are “short-lived”
- Can also be converted to RFC 3820 proxy cert

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

Or...


Welcome To The CILogon Service - Iceweasel
Iceweasel
Welcome To The CILogon Ser...
cilogon.org
https://cilogon.
Google
CILogon Service
Show Help
Select An Identity Provider:
University of Washington
University of Wisconsin-Madison
University of Wisconsin-Milwaukee
Verisign
Search:
Remember this selection:
Log On
By selecting "Log On", you agree to CILogon's privacy policy.
For questions about this site, please see the FAQs or send email to help @ cilogon.org.
Know your responsibilities for using the CILogon Service.
This material is based upon work supported by the National Science Foundation under grant number 0943633.
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.

Had a NSF S3I2 existed...

We would have asked:

- What is the arc for "grid" PKI and GSI?
- Do the "grid" and "web" SSO communities talk?
- How is MyProxy evolving?
- How is GridShib evolving?
- What will be process for accrediting LIGO CAs?
- Do we need a HSM card? Which vendors?

Had a NSF S3I2 existed...

We should have been told:

- Many communities moving away from user managed PKI
- Little interaction between “grid” and “web” SSO communities but it has started and you can find it here...
- MyProxy has strong support and solid development
- GridShib is not evolving anymore
- Pay attention to CILogon
- Active push to remove HSM card requirement
- “Here is a roadmap for deploying a CA that can later be accredited by TAGPMA”

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

SAML2 ECP for non-browser web resources (RESTful WS)

Had a NSF S3I2 existed...

We would have asked:

- What is the ecosystem for non-browser apps?

Had a NSF S3I2 existed...

We should have been told:

- Watch project moonshot closely (EAP, GSS-API, RADIUS)
- "Here is a tutorial on the SAML2 ECP profile"
- Watch Shibboleth proposed GSS-API/SASL with ECP closely

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

Had a NSF S3I2 existed...

We would have asked:

- What are other VOs doing about email lists?

We should have been told:

- Stop using mailman already!
- Take a close look at Sympa

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,CILogon, no grid access
5. No access to code repositories
6. No email lists

Had a NSF S3I2 existed...

We would have asked:

- Disabling the Kerberos principal is good, yes?

We should have been told:

- It's too good actually. Most VOs are going to want to slowly evaporate access based on the resource and the role of the user in the collaboration. You need to focus on authorization and access control more and less on authentication. Invest the time to understand how you want to offboard various user roles.

SAML federation is next step

Use cases:

- Collaboration with Virgo (France, Italy)
- Collaboration with LCGT (Japan)
- Astronomy community collaboration spaces
- CILogon
- Globus Online
- NSF program managers
- External advisory panel members
- Condor collaborators to help with troubleshooting
- ISI collaborators to help with troubleshooting
- Consuming federated identities within LIGO

SAML federation is next step

LIGO days(?) away from joining InCommon

Intend to also pursue international federations

- Virgo pursuing Fédération Éducation-Recherche and IDEM
- GakuNin (Japan)
- DFN, UK AMFER, Australian Access Fed,...

SAML federation is next step

LIGO Cybersecurity Officer “has concerns about federation”

Can we really trust those other people?

Had a NSF S3I2 existed...

We would have asked:

- Help!
- How do I enage with my security officer?
- How do I characterize change in risk profile due to SAML federation?

We should have been told:

- "Here is a document discussing the benefits and risks for science VOs when participating in SAML federations. It is intended to be consumed by both architects and security staff."

An Analysis of the Benefits and Risks to LIGO When Participating in Identity Federations

by Jim Basney, Scott Koranda, Von Welch

`https://dcc.ligo.org/public/0070/G1100964/002/LIGOIdentityFederationRiskAnalysis.pdf`

Had a NSF S3I2 existed...

It is difficult to estimate, but I expect if a NSF S3I2 had existed and offered non-biased consulting services around IdM and cybersecurity LIGO would have saved two years of senior FTE effort.*

*FTE effort many smaller VOs do not have