CTSC
CENTER FOR TRUSTWORTHY
SCIENTIFIC CYBERINFRASTRUCTURE

A Study of Three Approaches to International Identity Federation for the LIGO Project

LIGO-G1300686-v2

October 15, 2013

For Public Distribution

Jim Basney and Scott Koranda

About CTSC

The mission of the Center for Trustworthy Scientific Cyberinfrastructure (CTSC, trustedci.org) is to improve the cybersecurity of NSF science and engineering projects, while allowing those projects to focus on their science endeavors. This mission is accomplished through one-on-one engagements with projects to solve their specific problems; broaden education, outreach and training to raise the practice-of-security across the community; and looking for opportunities for improvement to bring in research to raise the state-of-practice.

Acknowledgements

The authors would like to thank Stuart Anderson, Warren Anderson, Steven Carmody, Scott Cantor, John Krienke, Maria Laura Mantovani, Chris Phillips, Hiroyuki Sato, Tom Scavo, Rhys Smith, Ian Young, and the LIGO Identity and Access Management project team.

This document is a product of the Center for Trustworthy Scientific Cyberinfrastructure (CTSC). CTSC is supported by the National Science Foundation under Grant Number OCI-1234408. For more information about the Center for Trustworthy Scientific Cyberinfrastructure please visit: http://trustedci.org/. 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.

Using & Citing this Work

Cite this work using the following information: J. Basney and S. Koranda, “A Study of Three Approaches to International Identity Federation for the LIGO Project,” Center for Trustworthy Scientific Cyberinfrastructure, trustedci.org, May 2013.

This work is available on the web at the following URL:
[http://hdl.handle.net/2022/16760]

1 Background

The Laser Interferometer Gravitational-Wave Observatory (LIGO) is a large research project funded by the National Science Foundation. LIGO seeks to make the first direct detection of gravitational waves, use them to explore the fundamental physics of gravity, and develop the emerging field of gravitational wave science as a tool of astronomical discovery. Through a cooperative agreement with NSF, the California Institute of Technology (Caltech) and the Massachusetts Institute of Technology (MIT) jointly operate the LIGO Laboratory and its two observatories, one in Hanford, WA and one in Livingston, LA. The LIGO Scientific Collaboration is the international group of researchers carrying out the science of the LIGO Observatories as well as that of the GEO600 detector in Hannover, Germany. Today LIGO is a worldwide collaboration with more than 1000 members from across five continents.

Over the past few years LIGO invested significantly to develop a SAML-based single sign-on infrastructure. LIGO operates a Shibboleth Identity Provider (IdP) and provisions a LIGO electronic identity (branded as an “albert.einstein@LIGO.ORG” identity) for each collaboration member. The collaboration operates more than 50 Shibboleth service providers (SPs) that host a wide spectrum of services including wikis, document catalogs, event databases, and data investigation tools.

LIGO has planned from the beginning to leverage federated identities to address two primary use cases. First, although LIGO provisions an electronic identity for each collaboration member, many members have a pre-existing federated identity that could in principle be used to access LIGO SPs. By reducing the number and scope of provisioned LIGO identities the collaboration can decrease the burden of having to operate an IdP and the associated help desk services needed to assist users in managing a LIGO electronic identity. Second, the full impact of LIGO science can only be realized with close collaboration between LIGO scientists and astronomers and astrophysicists from other projects. Federated identity helps streamline collaboration between LIGO scientists and other researchers by enabling easier access to resources without the need for provisioning LIGO identities to external collaborators.

To facilitate leveraging federated identity and begin pursuing interoperability LIGO has joined the InCommon identity federation in the United States. Through the InCommon identity federation LIGO has enabled federated access to its resources for a large number of researchers in the US. Because LIGO is an international collaboration, however, and the pool of possible external collaborators is global, far more work remains to federate with institutions and projects from around the world.

Today the only choice for LIGO is to negotiate peer-to-peer federation with each IdP and SP that supports its international collaborators. This approach, however, does not scale since the number of IdP and SP targets for international federation is so large and the negotiation of policy, for example regarding privacy and attribute release, so time consuming.

Fortunately most international institutions hosting LIGO members or collaborators are themselves part of an existing SAML identity federation and LIGO need not pursue interoperability with each individual institution or organization. Rather LIGO could join each existing identity federation of interest.

Still, the number of identity federations of interest to LIGO is large. Today the federations that either intersect directly with LIGO membership or with existing and possible LIGO collaborators includes:

- Australian Access Federation (AAF)
- FederationCAFe (Brazil)
- Canadian Access Federation (CAF)

- CERNET Authentication and Resource Sharing Infrastructure (China)
- DFN-AAI (Germany)
- Fédération Éducation-Recherche (France)
- eduID.hu (Hungary)
- INFLIBNET Access Management Federation (India)
- IDEM (Italy)
- GakuNin (Japan)
- SURFnet (Netherlands)
- Servidor de Identidad de RedIRIS (Spain)
- UK Access Management Federation for Education and Research

Of special interest are collaborators from other interferometric gravitational wave experiments and organizations including the European Gravitational Observatory (EGO), responsible for the computing and networking for the Virgo (French and Italian) interferometer experiment, and KAGRA (Japan), as well as the planned LIGO facility to be located in India in the near future. At this time, to facilitate research, LIGO provisions a LIGO electronic identity for Virgo members who request access to LIGO resources. LIGO would prefer to leverage federated identity instead, since provisioning and managing identities for Virgo members is burdensome. At this time, however, EGO project representatives have indicated that they do not expect to be ready to join the IDEM federation until 2015.

To realize the promise of federated identity to enable easier collaboration, LIGO is faced with the daunting task of pursuing interoperability with each of the identity federations separately, most likely by having to directly join each federation.

A better path to international federation for LIGO would be to leverage its existing membership in InCommon. Ideally, having already joined InCommon, LIGO would automatically interoperate with the federations listed above, as well as other identity federations throughout the world, through inter-federation agreements, policies, and practices. A vetted research and scholarship organization such as LIGO, after joining InCommon, should find that without further effort its IdP and SPs interoperate with any IdPs and SPs in any of the higher education and research SAML federations worldwide.

No functional inter-federation infrastructure between InCommon and other federations, however, exists today. InCommon has only recently begun negotiating with the UK Access Management Federation for Education and Research. The eduGAIN service in Europe is intended to enable federation between the GÉANT (GN3) partners' federations, but as this time no agreement between InCommon and eduGAIN has been announced.

2 Three Approaches to International Federation

Together CTSC and LIGO launched three simultaneous efforts to explore international SAML federation between LIGO and its collaborators. The three specific efforts were chosen to span the spectrum of federation approaches from point-to-point direct federation to bilateral federation agreements between existing large national SAML federations so that LIGO could better understand the policy and technical issues surrounding international federation, the timelines necessary for each, and begin to develop a long term strategy for international interfederation in support of LIGO's long term scientific mission.

Specifically the three efforts included:

1. Point-to-point federation between LIGO service providers and an IdP able to authenticate and assert attributes for members of the KAGRA project.
2. LIGO joining the Italian IDEM SAML federation operated by GARR, the Italian Research and Education Network (NREN),[1] in order to support federation between the LIGO service providers and IdPs able to authenticate and assert attributes for some members of the Virgo project.
3. A bilateral federation between InCommon in the US and the UK Access Management Federation for Education and Research (UK Federation) to leverage LIGO's existing investment in InCommon and support federation between LIGO service providers and IdPs able to authenticate and assert attributes for both members of LIGO at UK institutions and astronomy and astrophysics collaborators at UK institutions.

The point-to-point federation with a KAGRA IdP was underway already when the CTSC and LIGO engagement began but was continued and focused with CTSC effort. CTSC and LIGO staff initiated the other two efforts directly as part of the CTSC and LIGO engagement.

3 Peer-to-peer Federation with the KAGRA IdP

The LIGO and KAGRA federation effort began in November of 2010 when Scott Koranda from LIGO, Kazu Yamaji from GakuNIN, and Hiroyuki Sato from GakuNIN met at an InCommon members meeting and explored the idea of GakuNIN deploying and managing an IdP for the KAGRA project. Initial progress was slowed in part by sensitivities within the LIGO and KAGRA communities over plans to share information since formal agreements to collaborate had not yet been signed. Roughly one year later, in November of 2011 the first integration between a LIGO SP and the KAGRA IdP was in place to support a less important and not well focused use case.

In November of 2012 work began in earnest to federate the LIGO SP hosting the Document Control Center (DCC) in order to enable sharing of some LIGO instrument design documents with KAGRA. Then in January of 2013 the LIGO SP hosting the main LIGO wiki was federated with the KAGRA IdP to support joint committee work between LIGO and KAGRA. These two important federation use cases supporting direct LIGO and KAGRA collaboration helped drive the peer-to-peer federation work and the KAGRA IdP from a prototyping phase to full production and it is expected that the KAGRA IdP will continue to function in production going forward to support KAGRA.

Currently the KAGRA IdP supports 40 KAGRA users. Of those 40 users 23 are from the University of Tokyo and 17 are not. Since the IdP went into production it has averaged roughly 30 unique authentication events per month leading to the access of LIGO resources.

The peer-to-peer federation exercise explored and highlighted a number of federation and interoperability issues:

[1] https://www.idem.garr.it/en

SAML metadata exchange Because the project began as a prototype effort the initial exchange of metadata between the LIGO SPs and the KAGRA IdP was done through email with nothing more than coordinated bootstrap trust. The DCC federation initially involved two development or test servers before federation of the final production SP and each time metadata was exchanged “by hand” over email. The LIGO SPs consume a separate metadata feed prepared and signed by LIGO that contains the KAGRA IdP metadata.

More recently the KAGRA IdP has begun directly consuming a signed LIGO metadata feed that includes the LIGO SPs so that further federation of LIGO SPs with the KAGRA IdP will not require direct exchanges of metadata.

SAML UI metadata for KAGRA IdP To facilitate IdP discovery both the DCC and wiki SPs deploy the Shibboleth Embedded Discovery Service (EDS). The EDS user experience is best when the service harvests user interface metadata elements (mdui) to display a logo for each IdP. Because the KAGRA IdP began as a prototype effort and is not directly operated by KAGRA it was necessary for LIGO to obtain and host an appropriate KAGRA logo for use with the LIGO EDS deployments. Work is underway now to transition hosting of the KAGRA logo by the KAGRA IdP server.

Attribute assertion by KAGRA IdP The KAGRA IdP asserts eduPersonPrincipalName (ePPN), givenName, sn, and mail. The ePPNs have the form `uid@shibbi.pki.itc.u-tokyo.ac.jp`, though not all KAGRA users are from the University of Tokyo.

No specific policy negotiations between LIGO and the GakuNIN operators for the KAGRA IdP were necessary in order for the KAGRA IdP to assert the attributes listed above. This simple fact greatly simplified the peer-to-peer federation work.

Access control Since the KAGRA IdP does not assert any group, role, or entitlement information regarding the KAGRA users and no federated group information about the KAGRA community is available, all access control at the LIGO SPs is currently done using either name-based access control against the asserted ePPNs or by application-specific group and privilege information. Plans are underway for LIGO to deploy an instance of COmanage[2] along with an attribute authority to enable management of the KAGRA groups by KAGRA members and assertions and consumption of the group details by LIGO SPs.

We note this simple metric for the LIGO/KAGRA peer-to-peer federation exercise: over the course of three years approximately 180 email messages have been exchanged to coordinate the federation.

4 LIGO membership in the IDEM Federation

As part of the joint CTSC and LIGO engagement to explore international SAML federation LIGO decided to join the Italian IDEM federation. The choice of IDEM as opposed to one of the other national SAML federations listed above was influenced by:

- IDEM including a number of IdPs, especially from the Istituto Nazionale di Fisica Nucleare (INFN), that support members of the Virgo project.

[2] `https://spaces.internet2.edu/display/COmanage/Home`

- LIGO's desire to help expedite in any way it can the transition of the Virgo collaboration from using LIGO-issued credentials to access LIGO resources to leveraging a fully federated infrastructure.
- Scott Koranda's direct face-to-face interactions with Maria Laura Mantovani, a senior member of the GARR IDEM staff.

Organizations that are not members of GARR may still join IDEM as partners, as opposed to full members. The application process for partners only requires the signing of a Memorandum of Understanding (MOU) rather than a legal contract. Initially LIGO/CTSC staff anticipated that because a legal contract was not required to join IDEM and a MOU would suffice the application process would proceed quickly.

Due to some of the language in the MOU, however, LIGO Laboratory staff concluded it necessary to involve the Caltech legal department in the MOU process. This in turn has slowed the process and as of now the MOU forms have not been signed and returned to IDEM. We do not anticipate any problems that will prevent Caltech from signing the MOU on behalf of LIGO at this time but do not expect the transaction to be completed by the end of the LIGO/CTSC engagement.

5 Interfederation with the UK through InCommon

To leverage LIGO's existing investment in InCommon and through it pursue international federation the CTSC/LIGO staff chartered the InCommon Technical Advisory Committee (TAC) Interfederation Subcommittee.[3] The mission of the subcommitee

> "is to promote and pursue interfederation between the InCommon Federation and other SAML federations via a community-based process. The subcommittee makes recommendations to the InCommon Technical Advisory Committee, and members of the subcommittee interact with members and operators of other SAML federations to draft agreements and common practices. REFEDS is the preferred forum for cross-federation discussions, and InCommon-specific discussions take place on the subcommittee mailing list and on subcommittee phone calls. The subcommittee does not make agreements on behalf of InCommon or represent InCommon in any official capacity. Both policy and technical aspects of interfederation are in scope for the subcommittee."

Deliverables for the subcommittee are

1. Documentation of InCommon community interfederation use cases and timelines, including the international collaborations of the Laser Interferometer Gravitational-Wave Observatory (LIGO) project.
2. Documentation of plans and/or issues for interfederation with UK Access Management Federation, Australian Access Federation, Canadian Access Federation, eduGAIN, and other federations of interest to the InCommon community.
3. Documentation of lessons learned, recommendations, and potential future work areas/items for InCommon to consider on the topic of interfederation.

[3] https://spaces.internet2.edu/display/incinterfed/Interfederation+TAC+Subgroup

4. Work summary to TAC at end of work.

(While this document covers some of the same material it is not part of the subcommittee deliverables and we refer the reader to the full set of documentation at the subcommittee's web site.)

Members of the committee include the CTSC/LIGO engagement staff, InCommon Operations staff, UK Federation staff, and interested members of the broader community. The subcommittee is expected to complete all deliverables and either close or recharter by the end of June 2013.

To investigate and support the LIGO use case the committee pursued an exchange of select metadata between InCommon and the UK Federation. The effort built upon work already underway in the UK to support interfederation trials.[4]. The committee focused on the specific use case of federating a Cardiff University IdP with the LIGO SP supporting the main LIGO wiki, with the goal of allowing both LIGO collaboration members at Cardiff and their colleagues with interesting research interests in astronomy and astrophysics to reach the LIGO wiki using federated identities authenticated by the Cardiff IdP.

The LIGO SP metadata, as already registered in InCommon, was consumed and integrated with the UK Federation metadata feed already supporting the UK interfederation trials. Since InCommon had no similar trial underway at the time a new metadata feed was prepared by the committee[5] that combined the InCommon metadata and the UK Federation interfederation trial metadata into a new aggregate for consumption by the LIGO SP. The Shibboleth Metadata Aggregator is used for the aggregation. The aggregate is signed using a self-signed certificate used just for the purpose at this time.

After configuring the LIGO wiki SP to consume the aggregate that contains the UK Federation interfederation trials metadata with the test Cardiff IdP initial interoperability testing revealed configuration issues with the test Cardiff IdP. Most LIGO SPs will attempt to use SAML artifact resolution if the IdP advertises an artifact endpoint in the metadata. The Cardiff IdP does advertise and artifact endpoint but was not configured to support artifact resolution. More investigation showed this to be true for a number of IdPs in the UK Federation. The Cardiff IdP operator adjusted the configuration for the test IdP to properly support artifact resolution and after the adjustment more testing showed that Cardiff users successfully accessed the LIGO wiki using Cardiff identities.

The interoperability testing did expose, however, that further work is needed to be done by the LIGO wiki SP operator. The Cardiff IdP asserts opaque ePPN values in order to preserve the user's privacy. Preserving user privacy in this way is a not uncommon approach taken by IdP operators as a mechanism for complying with current UK law. It is expected that all IdPs operated by NRENs in Europe will, if they do assert ePPN at all, assert opaque values for users. Most often the IdPs will be expected to assert a completely opaque but persistent identifier for either ePPN or eduPersonTargetedID (ePTID).

The LIGO wiki SP at this time uses a simple algorithm to convert the asserted ePPN to a proper "wiki name" needed by the application (Foswiki version 1.1.5). For non-opaque values the conversion leads to a wiki name that is shown by the wiki as the person responsible for editing the page that users can consume, understand, and map to a colleague (eg. `skoranda@uwm.edu` is mapped to `SkorandaATuwmDOTedu` and the LIGO community knows that user as 'Scott Koranda'). When opaque values for ePPN are asserted by an IdP, however, the wiki name cannot be mapped to a colleague by users and that was the immediate reaction from the test users. The LIGO wiki SP operator recognizes that the application must be adjusted so that any value for ePPN or ePTID asserted can be mapped to

[4] `http://www.ukfederation.org.uk/content/Documents/InterfederationTrialFAQ`

[5] Special thanks to Steven Carmody from Brown University for his work to create and support the initial metadata aggregate feed, and to Ian Young from the UK Federation for his assistance.

a wiki name set by the users themselves and provisioned during an initial registration phase when the application is first accessed by the user. That work is scheduled.

Having demonstrated with the pilot project that LIGO is able to leverage an aggregated metadata feed that includes metadata for IdPs in the UK Federation, the LIGO staff with assistance from the subcommittee will formally petition the InCommon TAC and request that the metadata aggregate feed that combines InCommon metadata with that from the UK Federation be made a service deliverable from the InCommon Operations team. We expect the InCommon TAC to support the request and recommend to the InCommon Steering Committee that Operations take on the task with appropriate resource allocation as determined by Steering.

6 Observations, Conclusions, and Recommendations

At the end of the study only the KAGRA IdP, integrated with the LIGO SPs via a peer-to-peer federation, is being used in production to access LIGO resources. While it is true that the KAGRA federation exercise began long before the CTSC/LIGO engagement began, and the use cases for the KAGRA IdP have been more important recently to supporting LIGO and gravitational-wave science, this indicates that peer-to-peer federation agreements can play an important role for science organizations. While peer-to-peer federation arrangements do not scale well and circumvent a number of community best practices, they do expedite and enhance collaboration between scientists and ultimately that is the goal for applied identity management for scientific organizations.

The assertion of opaque values for ePPN by the Cardiff IdP highlight that SPs have little choice but to expect and plan to only receive opaque and most likely targeted identifiers for users. Efforts like the Research and Scholarship Category[6] in InCommon, while useful and worth supporting, are years away from being standardized and useful in an international context. SP operators wishing to federate with multiple international partners must plan to include either registration facilities at each SP for mapping from opaque identifiers asserted by the IdPs to useful user attributes, or must employ a centralized collaboration management tool, invitation service, or enrollment service along with an attribute authority that project SPs an query to retrieve not only project attributes about users but also the basic identifiers for users like given name, family name, and email. LIGO plans to deploy COmanage along with a Shibboleth attribute authority for this reason.

Since many web applications used by science projects and protected by a SAML SP are simple open source projects with few or no mechanisms for consuming identity, much less opaque identifiers that must then be mapped to the identifiers the application needs, much work remains to be done by science projects to prepare their web resources to consume federated identities.

We note that although the formation of the InCommon Interfederation Subcommittee was advertised to all InCommon participants, the only research organization that participated was LIGO. It is unlikely that of the 28 InCommon participants classified as government and nonprofit laboratories, research centers, and agencies only LIGO has international collaboration use cases that would benefit from interfederation. The InCommon efforts at interfederation would benefit from learning of the use cases and needs of the other organizations such as the Long Term Ecological Research Network. Involvement in particular by organizations like the Open Science Grid and XSEDE that can represent multiple science projects would be especially helpful, even while those organizations are still understanding how federated identity can be leveraged to further their community's efforts. In Europe the Federated Identity

[6] https://spaces.internet2.edu/display/InCFederation/Research+and+Scholarship+Category

Management for Research Collaborations[7] (FIM4R) group has gathered requirements across multiple science projects and documented the requirements in a report.[8] A similar requirements gathering effort in the US across science projects as input to InCommon would be valuable.

Until efforts like eduGAIN provide for ubiquitous federation between IdPs and SPs across international borders we expect larger science projects like LIGO to join multiple national identity federations. To facilitate non-legal entities like LIGO joining federations, we recommend that national identity federations develop specific policies and processes that do not require legal contracts or indemnification, and that do not use particular legal language likely to force projects to seek legal guidance. A standard application for research and scholarship service providers or organizations across national identity federations could greatly simply the process of federation for science projects.

[7] http://indico.psi.ch/conferenceDisplay.py?ovw=True&confId=2230

[8] http://cds.cern.ch/record/1442597