An Analysis of the Benefits and Risks to LIGO When Participating in Identity Federations[1]

Jim Basney[2], Scott Koranda[3], Von Welch[4]

[1] LIGO document number LIGO-G1100964-v2

[2] Senior research scientist at the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign where he leads the CILogon project. *jbasney@ncsa.uiuc.edu*

[3] Senior scientist in the LIGO group at the University of Wisconsin-Milwaukee and lead architect of the LIGO Identity Management project. *scott.koranda@ligo.org*

[4] Deputy Director of the Center for Applied Cybersecurity Research at Indiana University. *vwelch@indiana.edu*

About This Document

The LIGO Identity Management (IdM) project provisions and manages electronic identities for over 800 members that enable secure access to LIGO resources including data, the LIGO instruments, data analysis facilities, and general computing. To facilitate interactions and meet certain time demands the LIGO IdM project has also provisioned and manages electronic identities for groups of people not part of the LIGO project including collaborators from the Virgo project, NSF program managers, and advisory committee members. Each of these external collaborators receives a LIGO identity including a “login” and associated password.

As the nascent field of gravitational wave astronomy matures and more experimental gravitational wave facilities come online LIGO scientists will need to efficiently and securely collaborate with ever more experimentalists, astronomers, and astrophysicists from a multitude of research projects across disciplines and throughout the world. Secure and managed access to LIGO resources requires each individual to have an electronic identity. The LIGO IdM infrastructure then is confronted with a choice; it can either provision and manage an electronic identity for every individual seeking to access LIGO resources and collaborate and interact with LIGO project members, or it can leverage existing identities external to LIGO and already available to prospective collaborators. These electronic identities, being exchanged across security domains, are known as federated identities.

If the LIGO IdM infrastructure must provision and manage an electronic identity for every external collaborator the scale, complexity, and required funding for the IdM project must grow accordingly so that online collaboration is facilitated rather than frustrated. Since LIGO will not manage or control all web sites and other electronic resources needed to enable cross project and cross disciplinary collaborations, and it is unlikely that every other project will agree to accept LIGO identities without a reciprocal agreement, it will also be necessary for LIGO scientists to obtain and manage themselves more electronic identities provisioned by the other projects. Each LIGO scientist will be responsible for managing a handful of electronic identities needed to navigate throughout the day as she interacts with colleagues from other scientific projects.

Alternatively the LIGO IdM project can leverage existing identities its collaborators already posses and manage. Universities, colleges, institutes, laboratories, governments, and other organizations these collaborators are affiliated with have already issued and actively manage electronic identities for them. Many of these organizations have banded together to form identity federations. These federations streamline the technical work needed to leverage federated electronic identities and enable the federation members to collectively analyze, manage, and audit the benefits and risks of federation specifically and identity management more generally. By leveraging federated identities LIGO can focus more effort and money on tools, services, and resources for enhancing the efficiency of online collaboration and less on managing identities themselves.

Clearly federation can benefit the LIGO IdM project, LIGO scientists, and LIGO collaborators by allowing individuals to reuse and focus on a single, well managed and secure electronic identity. Federating is not, however, without risk and cost since interacting with and leveraging non-LIGO provisioned identities requires establishing and managing trust relationships between LIGO protected resources and the non-LIGO identity providers (IdP).

This document analyzes the benefits, costs, and risks to the LIGO project when federating with external organizations, including identity federations, and consuming federated electronic identities.

Document Scope

A wide variety of federated identity technologies and organizations exist that seek to form trust amongst organizations for online collaboration. This document focuses specifically on Security Assertion Markup Language (SAML) as a federation technology, the InCommon federation in the United States (US), and the benefits and risks to the LIGO project of joining the InCommon federation and leveraging federated identity to support its science mission. Although we focus on SAML and InCommon and analyze benefits and risks to LIGO in that context we do highlight where appropriate other technologies and organizations expected to intersect with LIGO due to its stature as an international project expected to significantly impact astronomy and astrophysics research for the coming decades.

Authorization Distinct from Authentication

We focus on trust management around federated identity and securely and authoritatively asserting a federated identity across security domains and do not address issues of authorization to resources. *We assume the reader understands that authentication by itself does not enable authorization to a resource, and that trusted identity assertions are only the first step in making trusted authorization decisions for access.*

The LIGO Project

LIGO is an ambitious project funded by the U.S. National Science Foundation (NSF) to detect gravitational waves and use them to observe and study astrophysical phenomena. The LIGO Laboratory is charged with building and operating interferometer detectors at the two LIGO sites in Hanford, WA and Livingston, LA and is operated jointly by the California Institute of Technology (Caltech) and the Massachusetts Institute of Technology (MIT) under a cooperative agreement with the NSF. The LIGO Scientific Collaboration (LSC) is an international collaboration of scientists and other researchers formed to contribute to the success of the LIGO science mission. Groups join the LSC by signing a memorandum of understanding (MOU) with the LIGO Laboratory. Especially when interacting with external entities the term "LIGO" is used to refer to the union of the broader LSC group and the LIGO Laboratory and we follow that convention in this document.

LIGO Identity Management Status and Plans

Since 2007 the LIGO IdM Project (formerly known internally to LIGO as the "Auth Project") has actively provisioned and managed individual electronic identities for all LIGO collaboration members. The identities, as a consequence of the deployment details, have come to be branded as "@LIGO.ORG" or "albert.einstein" identities since each provisioned identity is of the form "given.family@LIGO.ORG". These identities and the LIGO IdM infrastructure enable secure and efficient access to LIGO resources including web pages, web services, data analysis computing facilities, LIGO interferometer data, as well as some instrument controls. As noted above @LIGO.ORG identities have also been provisioned for Virgo members needing access to LIGO resources, NSF program managers, and various advisory panel members. Today over 1200 @LIGO.ORG identities have been provisioned and are actively managed by users and the LIGO IdM project.

Each @LIGO.ORG identity has associated with it a Kerberos principal and the IdM project manages a Kerberos KDC and distributed replicas. Users manage identity attributes including names, addresses, and telephone numbers as well as the password associated with the Kerberos principal via a set of web pages collectively known as "MyLIGO".

At this time the password associated with each Kerberos principal is not required to meet any specific policy. The LIGO Security Committee (SecComm) is actively working to configure the KDC to check passwords against standard password dictionaries, and has adopted a plan to require passwords with a minimum level of entropy meeting the FICAM level of assurance (LOA) 2 standard[5] (LIGO's current plan is to require a minimum of 15 characters). Users may request a password reset or change the password themselves via the MyLIGO web interface.

Authorization to LIGO resources is primarily group based. Group membership is recorded using Grouper and managed using a combination of MyLIGO, the Grouper UI, LDAP, and associated scripts and tools. LIGO resources consume group membership directly or indirectly from Grouper and LDAP and use that information to make authorization decisions and grant or deny @LIGO.ORG identities access to resources.

Users may use their @LIGO.ORG Kerberos principal to directly access Kerberos-enabled services including login via enabled OpenSSH servers and PAM-enabled consoles, some mail servers, and various code repositories using either HTTPS for transport or tunneled through OpenSSH.

LIGO protected web pages and web services use the Shibboleth[6] implementation of SAML2 to manage access. Users authenticate to the Shibboleth IdP using their Kerberos principal and

[5] http://www.idmanagement.gov/documents/TrustFrameworkProviderAdoptionProcess.pdf

[6] https://wiki.shibboleth.net/confluence/display/SHIB2/Home

password. The IdP asserts attributes about the LIGO electronic identities including group memberships that are later used for authorization decisions by the Shibboleth Service Provider (SP). The infrastructure is configured to support single sign-on across the LIGO web space. Some legacy LIGO web sites are not part of the Shibboleth infrastructure and use Kerberos directly for authentication.

Since 2001, and well before the LIGO IdM project began, LIGO participated in a number of national and international grid computing projects including GriPhyN, iVDGL, Grid3+, and the Open Science Grid (OSG). To facilitate access both to its own computing resources, collectively known as the LIGO Data Grid (LDG), as well as other resources made available by those projects, LIGO chose to federate with the U.S. Department of Energy's DOEGrids Certificate Authority (CA), as well as a number of other international CAs that are part of the International Grid Trust Federation (IGTF).

At this time a user’s X.509 digital certificate, used as a grid identity, is not tied directly to the user’s “albert.einstein” LIGO identity. The request, verification, and delivery of the X.509 identity occurs along a parallel but distinct path to that used for managing the LIGO identity. In the future the LIGO IdM Project plans to deploy and support infrastructure that enables the LIGO identity to be used to obtain a short-lived X.509 digital credential for use on the LIGO Data Grid and other grids so that users no longer need to request, retrieve, renew, and manage directly their own X.509 credential. The exchange of a LIGO identity for the short-lived certificate will happen using a Kerberos ticket and a MyProxy server configured to act as a short-lived certificate service or by accessing the CILogon web service and using a SAML assertion to request and receive a short-lived X.509 certificate.

SAML and Federated Identity Basics

SAML

SAML[7] is an XML-based open standard for exchanging authentication and authorization data between between an identity provider or IdP (a producer of assertions) and a service provider (a consumer of assertions) or SP, often a web server protecting some resource such as static web pages or a wiki. Although an IdP and a collection of SPs can exist in the same security domain, as is currently the case for the LIGO IdP and available SPs, the full power of SAML is seen in federations containing multiple IdPs and SPs, allowing users from multiple organizations to collaboratively access SP services.

SAML assumes the user has enrolled and is able to authenticate against or with at least one IdP. SAML does not, however, specify the implementation of the authentication service and each IdP

[7] http://www.oasis-open.org/committees/download.php/27819/sstc-saml-tech-overview-2.0-cd-02.pdf

deployer is free to choose an appropriate authentication service. SAML does provide for details about the authentication event including the level of assurance to be communicated as part of the assertion so that individual service providers may include that assertion as part of an access control decision. SPs rely on the IdP(s) to identify the user or principal and at the principal's request, the IdP securely passes a SAML assertion to the service provider. LIGO's production SAML infrastructure leverages SAML 2.0 which became an OASIS standard in March 2005.

Details about the form of SAML assertions, protocols, binding, and profiles are explained in appendix A.

Shibboleth SAML Implementation

The Shibboleth Project provides free, open-source implementations of both a SAML 2 enabled identity provider and service provider with a focus on supporting federation--the secure access of resources across security domains. Details about the SAML protocols that the Shibboleth IdP and SP support at this time are available in appendix B.

Name Identifiers

The name identifier <NameID> in SAML 2 and leveraged by Shibboleth is used to identify the person that the IdP has issued an assertion about. Name identifiers can be anything and an email address or a Kerberos principal name are common examples. LIGO uses the Kerberos principal as the SAML2 name identifier.

A name identifier must be properly scoped and reversible. Scope means that the given identifier is issued from and only makes sense within a given security domain. A name identifier without a scope is meaningless, because it could have come from any IdP in the world. For example, asserting just the name identifier "patrick.brady" is meaningless, but saying "patrick.brady" from "https://www.uwm.edu" provides the requisite scope and makes the identifier meaningful. Reversible means that the IdP can translate the identifier back into the precise user within the lifetime of the identifier.

More details about name identifiers are available in appendix B.

SAML Metadata

"Metadata" in this context refers to the configuration data consumed by an IdP or SP to enable the communication between them. Typically the metadata is an XML document (often digitally signed) consumed directly the Shibboleth IdP or SP that enumerates the set of trusted partners or relying parties and configures each how to communicate securely with the other (for example, by including public keys and URLs). The Shibboleth IdP consumes a metadata XML document by looking for entities in the document that act in SP roles while the SP consumes metadata looking for entities in the document that act in IdP roles. As such it is common for both IdP and SP entities to be described in the same metadata XML document. An entity in the SAML metadata document is simply a server acting in the role of an IdP or SP. Each entity is required to have a unique name or entityID that distinguishes it from any other. Care must be taken such that each entityID is unique across the collective metadata for any identity federation. This check is usually managed by the appropriate federation administrators responsible for the management of the metadata and typically relies on proof of ownership of the DNS name in the entityID.

The Shibboleth software supports consuming both locally maintained metadata (usually a file or files on a file system) and remotely maintained metadata (usually from a published URL). Both locally and remotely maintained metadata can be digitally signed and the Shibboleth software configured to check and verify the signature using a well-known key or certificate before consuming the metadata.

Trust Model and Trust Management

The Shibboleth software implicitly trusts the metadata it has been configured to consume (provided the metadata, if signed, has been properly verified). Details on how the Shibboleth software implements and manages the trust are available in appendix C.

Identity Federations

Federated identity management involves having a common set of policies, practices and protocols in place to manage the identity and trust of users and services across organizations and security domains. Typically identity federations define a trust fabric, provide a set of agreed-on attributes used for exchanging information, offer software to enable authentication and authorization, and distribute the metadata necessary for interoperability.

The InCommon Federation

The InCommon Federation supports a framework for trustworthy shared management of access to on-line resources in support of education and research in the United States. InCommon provides a framework of shared policies, trust-establishing processes, and technology standards for IdPs and SPs to follow to streamline collaboration with multiple organizations. Rather than spend time establishing operating principles, technology hooks, and agreed-upon data exchange elements with each prospective partner, InCommon participants accomplish these once through InCommon and then leverage these common elements for many relationships.

In its own words:

> *“InCommon is a formal federation of organizations focused on creating a common framework for collaborative trust in support of research and education. InCommon makes sharing protected online resources easier, safer, and more scalable in our age of digital resources and services. Leveraging SAML-based authentication and authorization systems, InCommon enables cost-effective, privacy-preserving collaboration among*

> *InCommon participants. InCommon eliminates the need for researchers, students, and educators to maintain multiple, password-protected accounts. The InCommon federation supports user access to protected resources by allowing organizations to make access decisions to resources based on a user's status and privileges as presented by the user's home organization.”*[8]

Structure and Overlap with LIGO Scientific Collaboration

The structure of InCommon and its overlap with institutions that have an active LIGO Scientific Collaboration group are detailed in appendices D and E.

Federation Operating Policies and Practices

InCommon publishes a Federation Operating Policies and Practices (FOPP) document[9] describing at a high level the structure and operation of the federation in accordance with the LLC Agreement of InCommon ("Company Agreement") and Bylaws of the InCommon LLC. All InCommon participants and prospective participants are asked to review the FOPP to help assess what potential risks, if any, might be incurred by participation in the federation and to evaluate the level of assurance of the federation's services to ensure trustworthy operations and determine whether they meet a participant's minimum requirements. The complete evaluation of the entire federation's infrastructure and level of assurance is out of scope for the FOPP (since it would need to include evaluation of all relevant participants' policies and practices), and reviewers are asked to contact the InCommon office for clarification or additional information regarding the FOPP or other federation matters.

Participant Operational Practices

Each InCommon participant describes in a participant operational practice (POP) document its identity management system(s).[10] SPs can then use the POP to determine their level of trust for assertions from each participant IdP. Each IdP can evaluate each SP’s privacy policies, attribute collection, and use policies. It is a federation requirement that POP statements be publicly posted (but not necessarily advertised) on a website (the URLs for InCommon participant POPs are available to all active InCommon participant administrators via the secure administrative interface).

Some example POP documents can be found at the following URLs:

Institution	URL
Cornell University	http://www2.cit.cornell.edu/services/identity/InCommon.html
Louisiana State University	http://itsweb.lsu.edu/ITS_Security/files/item3676.pdf

[8] http://incommon.org

[9] http://www.incommon.org/docs/policies/InCommonFOPP_v20071015_Final.pdf

[10] http://www.incommon.org/docs/policies/incommonpop_20080208.pdf

Penn State	http://ait.its.psu.edu/services/identity-access-management/identity/accounts/pop/
University of California, Berkeley	https://calnet.berkeley.edu/idm/InCommon_POP_UCB.pdf
University of Massachusetts Amherst	http://www.oit.umass.edu/sites/oit.umass.edu/files/2011/07/11/incommon.pdf
University of Michigan	http://www.itcs.umich.edu/identity/incommon/pop.php
University of Wisconsin-Madison	http://www.doit.wisc.edu/middleware/InCommonPOP.pdf

Federation participants have reported that the process of preparing a POP has helped them identify weaknesses in the local identity management system and processes and provided a natural on-ramp for developing and putting into practice formal auditing procedures. There are, however, risks associated with POP documents because of the natural tendency to be speculative and document how things should work instead of how they actually work. POP documents can become stale and there is no formal requirement that they be audited.[11]

To address these shortcomings of the POP document InCommon has begun its Identity Assurance Program[12] and introduced the Identity Assurance Assessment Framework[13]. The framework describes the process by which an IdP becomes certified by InCommon as compliant with an Identity Assurance Profile (IAP), including the assessment and audit process and the specific qualifications an auditor must have in order to perform the assessment. The deliverable from the assessment process is an audit report to the IdP operator and a summary of findings report delivered to InCommon. Using the audit report InCommon determines whether one or more Identity Assurance Qualifiers can be used by the IdP. Once approved by InCommon, the IdP may then include the appropriate Identity Assurance Qualifier(s) as part of the assertions it makes.

At this time InCommon has defined the Bronze and Silver Identity Assurance Profiles (IAP). [14] The profiles are intended to be compatible with the US federal government ICAM Trust Framework Provider Adoption Process, Levels of Assurance 1 and 2[15]. The Bronze profile

[11] https://spaces.internet2.edu/display/CAMPJune2011/CAMP+Resources

[12] http://www.incommon.org/assurance/

[13] http://www.incommon.org/docs/assurance/IAAF_V1.1.pdf

[14] http://www.incommon.org/docs/assurance/IAP_V1.1.pdf

[15] http://www.idmanagement.gov/documents/TrustFrameworkProviderAdoptionProcess.pdf

focuses on sequential identity and the reasonable assurance that the same person is authenticating each time with a particular credential. While no identity proofing requirements are specified, it is expected that IdPs use reasonable care when issuing credentials to confirm that a single individual applies for and receives a given credential and its authentication secret. Bronze qualified assertions are typically usable by individuals seeking access to online information resources licensed to an organization and for which the subject is an eligible user. They are also usable for access to services where the SP invokes other methods for linking of the subject identifier to information the SP already has regarding individuals who should have access to its services.

The Silver assurance profile builds on the Bronze profile requirements and adds criteria for individual identity proofing and identity information records. Stronger credential technology(ies) and credential management are also required. The Silver profile is intended to assure a reasonably strong binding between the physical subject and that subject's digital credential. Credentials must at a minimum make use of authentication secrets that are sufficiently difficult to guess or intercept.

A table summarizing all of the identity assurance criteria defined for Bronze and Silver IAPs is shown in appendix F.

At this time InCommon is preparing to accept the first round of auditor reports and to approve the first set of IdPs that may include the appropriate Identity Assurance Qualifier(s) as part of the assertions they make.

InCommon SAML Metadata

The InCommon SAML metadata is the basis for trust within the federation. The trust model or trust management uses the "inline" model as detailed above, as opposed to a more traditional X.509 certificate-based PKI. Federation participants trust InCommon to vet the metadata content submitted by other participants, and InCommon vouches for the integrity of the metadata it makes available to participants.

The InCommon Federation metadata is published in the following location: http://wayf.incommonfederation.org/InCommon/InCommon-metadata.xml. The metadata is digitally signed, and InCommon strongly recommends that participants refresh metadata daily to ensure the most up-to-date keys and other registered information are available to all SPs and IdPs. InCommon also publishes a diff of consecutive metadata files every time a new metadata document is published and the diff is sent to an e-mail list for convenience of review.

Participants submit metadata to InCommon through the administrative web interface. InCommon staff processes metadata submissions within one (1) business day, except during Internet2 Member Meetings (week long meetings held twice a year) when delays are expected. Typically, submissions are reviewed Monday through Friday, at approximately 2:30pm Eastern Time, and published at approximately 3:00pm Eastern Time, although exact times may vary.

InCommon Metadata, Shibboleth Mechanics, and Authorization

We emphasize that joining InCommon and gaining access to the InCommon SAML metadata does not force LIGO to accept any external identities from any IdP, nor does it force LIGO to interoperate with any SP. Gaining access to the metadata allows LIGO, when it is ready, to add the metadata for its own IdP if LIGO wishes to enable its users to use LIGO identities to access non-LIGO resources. LIGO may also when it is ready choose to add the metadata for any specific LIGO SP that it wishes to interoperate with non-LIGO IdPs.

Before a LIGO SP will interoperate with any IdP the SP must be configured with the metadata for that specific IdP. We expect the LIGO IdM project to centrally consume the InCommon metadata, verify it, filter it to pick out only the IdPs with which LIGO is willing to federate, and then add that metadata to the centralized and signed metadata that LIGO distributes to its SPs (the SPs use a centrally managed configuration to pull signed metadata issued by LIGO). This is a standard approach taken by many InCommon participants. Likewise we expect the LIGO IdM project to consume the InCommon metadata and filter to pick out only the non-LIGO SPs with which the LIGO IdP is willing to interoperate. The filtering may be granular and only include specific SPs--there is no requirement that all the SPs from an organization be included. After federating the LIGO IdM project absolutely still controls the implementation of the policy deciding with which external entities LIGO will interoperate and does not need to delegate that decision to the SP operators. Note also that if necessary per-SP metadata configurations can be implemented in a straightforward way.

Lastly we want to emphasize again that interoperating with an external IdP does not imply that every user known to the IdP has access to LIGO resources--LIGO still controls the authorization to all resources.

International and Other Federations Relevant to LIGO

REFEDS (Research and Education Federations) mission “is to be the voice that articulates the mutual needs of research and education identity federations worldwide. It aims to represent the requirements of research and education in the ever-growing space of access and identity management”.[16] At the REFEDS wiki one can find a survey of federations from around the globe.[17] Below we articulate a list of identity federations that are of particular interest to LIGO (all federations listed use a SAML infrastructure):

- Deutches Forschungsnets (DFN) runs a production federation called DNF-AAI available to research and higher education institutions in Germany.
- The UK Access Management Federation for Education and Research runs a production federation available to all research and educational institutions in the UK.

[16] https://refeds.org/

[17] https://refeds.terena.org/index.php/Federations

- The Australian Access Federation runs a production federation available to all Australian research and educational institutions in Australia and New Zealand.
- Fédération Éducation-Recherche and IDEM provide federated identity services available to research and higher education organizations in France and Italy respectively. LIGO collaborators in the Virgo project have voiced plans to interoperate with those federations, providing a path that would enable LIGO to stop issuing and managing electronic identities for Virgo collaboration members.
- GakuNin is the identity federation for research and educational institutions in Japan.

Anticipated LIGO Use Cases for Identity Federation

Below we detail some anticipated LIGO use cases for identity federation. Not all of the use cases would be directly facilitated by LIGO joining the InCommon federation but we mention them to give a fuller picture of the possible federation use cases. We group anticipated LIGO use cases for identity federation into two different scenarios--federation where the LIGO IdP authenticates users who then access SPs not managed by LIGO, and federation where a non-LIGO IdP authenticates users who then access LIGO SPs.

Federation starting with the LIGO IdP

- **Accessing Virgo resources**: Currently LIGO users needing access to Virgo resources must be given a Virgo identity. The Virgo infrastructure is based on Active Directory Federation Services (ADFS), which can be federated with a Shibboleth SAML infrastructure. Detailed federation instructions specific to InCommon are available.[18] After proper federation LIGO users could access Virgo resources using LIGO identities.
- **Accessing LCGT resources**: At this time there is no formal relationship between LIGO and LCGT requiring access to resources within a project by members from the other project. We expect, however, that as LCGT matures scientists from both projects will need to access resources managed by the other project in order to enable joint work similar to the ongoing work between LIGO and Virgo. Managers from the GakuNin federation expect to work with the LCGT project to leverage federated identities and have begun collaborating with managers from the LIGO IdM project to plan for federation between LIGO and GakuNin to allow LIGO scientists to use their LIGO identities to access LCGT resources.
- **Accessing collaboration spaces**: We expect that in the Advanced LIGO era LIGO scientists will routinely need to access shared collaboration spaces (wikis, web pages, data portals) with astronomers, astrophysicists, and numerical relativists. Rather than needing to obtain unique credentials for each resource LIGO scientists could use their federated LIGO identity to access the resources.
- **Research.gov**: The National Science Foundation (NSF), through Research.gov, enables federated access to NSF Fastlane for InCommon participants. We expect LIGO

[18] https://wiki.shibboleth.net/confluence/download/attachments/4358293/ADFS_and_Shib.pdf

scientists, especially those from institutions that are not part of InCommon, would find federated access to Fastlane using LIGO identities convenient.

- **CILogon**: The CILogon Service[19] allows users to authenticate with federated identities and obtain X.509 certificates for secure access to cyberInfrastructure such as XSEDE and OSG. More information about using federated identities for access to cyberinfrastructure, including the role that the CILogon Service plays, is provided in the *Roadmap for Using NSF Cyberinfrastructure with InCommon*.[20] We expect LIGO users would find it convenient to use their federated identities to obtain short-lived X.509 credentials from CILogon instead of having to manage their own X.509 certificate and key, and use the short-lived credential to access not only XSEDE and OSG but also the LIGO Data Grid and European grid resources.
- **Globus Online**: Globus Online[21] makes robust file transfer capabilities accessible to any researcher with an Internet connection and a laptop. Users schedule transfers by authenticating to the Globus Online web portal, browsing “endpoints”, and using a “click and drag” approach. The Globus Online team is pursuing membership in InCommon and has agreed to work with LIGO to configure profiles for LIGO users with endpoints for the LIGO Data Grid pre-defined.

Federation starting with a non-LIGO IdP

- **Virgo and LCGT**: As noted above we expect significant collaboration between Virgo, LCGT, and LIGO scientists using federated identities. In this use case Virgo and LCGT scientists could use non-LIGO federated identities to access LIGO resources, rather than requiring that they maintain separate LIGO identities.
- **PAC and other external advisory panel members**: The members of various external panels that advise LIGO throughout the year could access the web pages and wikis used during the advisory process with federated identities.
- **NSF program managers:** The NSF has joined InCommon and can consume federated identities but at this time it does not run a production IdP. When the NSF does manage a production IdP then both the LIGO, Gravitational Physics, and any other relevant NSF program managers could use their federated identities to access LIGO resources.
- **ISI collaborators**: LIGO and the Pegasus team at the Information Sciences Institute (ISI) at the University of Southern California (USC) have enjoyed a long collaboration. A number of ISI staff have been given LIGO identities in order to access LIGO resources and assist data analysts with managing analysis workflows. Since USC is an InCommon participant they could instead use their federated identities to access LIGO resources.
- **Condor collaborators**: LIGO and the Condor team at the University of Wisconsin-Madison (UWisc) Department of Computer Science have also enjoyed a long collaboration. Since UWisc is an InCommon participant we anticipate enabling access to

[19] http://www.cilogon.org/

[20] http://www.incommon.org/cyberroadmap.html

[21] https://www.globusonline.org/

LIGO resources for certain members of the Condor team in order to further facilitate collaboration, especially joint troubleshooting and technical support.

- **LIGO Open Data**: Federated identities are expected to play an important role in the era of open access to LIGO data, enabling collaboration and profile management for many in the astrophysics, astronomy, and the general public.
- **Internet2 collaborators**: The LIGO Identity Management Project (Auth Project) relies heavily on Internet2 middleware including Shibboleth, Grouper, and soon COmanage. Further collaboration between LIGO and Internet2 staff will be enhanced when appropriate Internet2 staff can access certain LIGO resources using their existing federated identities.
- **LSC members**: As detailed in appendix E there is a large intersection between InCommon participants and institutions with active LIGO Scientific Collaboration (LSC) groups. Further, many of the LSC groups outside of the United States have access to federated identities through national federations in Europe, Australia, and Asia. Some LSC members may find it convenient to use their federated "campus" identity as their single "working" or professional identity and access LIGO resources using it.

  Note that an informal survey of LSC principal investigators (PIs) from campuses that are InCommon participants and members of the Committee on Institutional Cooperation (CIC) found that the PIs were actively interested in using a single campus federated identity as their professional identity and to access LIGO resources.[22]

Benefits to LIGO of Joining InCommon

We expect LIGO to realize the following benefits from joining identity federations including InCommon:

- **Better experience for users**: Users, both in and outside of the LIGO collaboration, do not have to manage an array of accounts and passwords in order to participate in the gravitational wave experimental and astronomy communities. Single sign-on allows users to access any number of resources while signing on only once, and users only need to trained to securely manage a single identity.
- **More collaboration**: With the widespread availability, especially in other countries, of federated identities, LIGO can expect to leverage economies of scale by reducing or removing the need to repeat integration work for each new collaborative effort between LIGO and major scientific communities including the astronomy, astrophysics, and numerical relativity communities. By lowering the threshold for the user and making it easier for users to access collaborative spaces we expect to enable more scientific collaboration.
- **Reduced account overhead**: The creation and management of LIGO identities,

[22] Private communication with Scott Koranda as preparation for a talk at a CIC meeting on the CIC role in enabling the use of federated identity in university research projects.

especially for those people that are not part of the collaboration such as NSF program managers, PAC and other external advisory committee members, and collaborators from the astrophysics, astronomy, and numerical relatively communities, can be significantly reduced, requiring less effort from the help desk and system administrator staff.

- **Faster and more accurate integration and provisioning for new users**: After the appropriate manager, such as an LSC group PI, has authorized a federated identity asserted by a known and trusted IdP to join the collaboration, the integration and provisioning of that identity into the LIGO infrastructure can happen faster and more accurately without the need to provision and manage a new LIGO identity for the user. Working with known and trusted IdPs enables immediate assertions of many of the necessary attributes needed for provisioning without having to have the PI or user fill out (possibly mistakenly) a form to gather yet again the same attributes.
- **Economies of scale for contractual agreements**: Some or all of the policy and legal requirements for bilateral agreements between LIGO and certain classes of resource providers, such as Globus Online, may be consolidated by or leveraged from the federation policies, agreements and requirements documents, requiring less effort for LIGO to leverage those services. Note that Internet2 and InCommon are together actively pursuing agreements with cloud vendors for both computing and data storage that might be of interest to LIGO.

Changes to LIGO Risk Posture from SAML Federation

Today LIGO IdM is based on trust in a set of services LIGO operates itself such as the @LIGO.ORG kerberos realm, a set of PKI services run by Department of Energy's DOE Grids Certificate Authority, and the users themselves (to properly manage passwords, the private keys for X.509 certificates, kerberos tickets, and RFC 3820 proxy certificates). The primary change to LIGO's current risk posture at a high level is the addition of the non-LIGO IdPs into this mix as a trusted party responsible for provisioning and managing users' passwords at their respective institutions.

High Level Discussion of Changes to Risk Posture

Joining a SAML federation increases the risk for LIGO whether federated identities are consumed by LIGO services or LIGO identities are used to access non-LIGO services. Below we detail in a general way how LIGO's risk exposure is increased or changed:

- **Trust non-LIGO IdP operators**: LIGO must analyze each of the POP documents and decide to trust specific IdP operators that they will adequately manage the IdP such that LIGO can be reasonably assured to the required level of assurance that the assertions from the IdP can be trusted. LIGO must trust that the IdP operators and institutions properly implement and manage the necessary credential policies as outlined in the POP. Note that joining a federation like InCommon does not imply that LIGO would trust all IdPs in InCommon, but rather LIGO can choose to trust specific IdPs either across

LIGO or on a per SP basis. Trusting non-LIGO IdP operators, however carefully, still does increase LIGO's risk.

- **Trust federation staff to manage metadata**: The technical implementation of the trust management requires that LIGO trust federation (InCommon) staff to properly and adequately vet and manage the InCommon SAML metadata since as detailed above it is that metadata that enables a LIGO SP or IdP to consume SAML assertions and make the binary choice of whether to accept and consume or reject an identity assertion. As noted above, we expect the LIGO IdP project to centrally manage and appropriately filter the federation metadata so that only IdPs allowed by LIGO policy are made known to the LIGO SPs. Still, the correctness and veracity of the metadata relies on the infrastructure that the federation has deployed to ingest and then distribute the metadata and on the staff doing the work. Trusting the federation infrastructure and staff increases LIGO's risk.
- **Trust non-LIGO SPs to responsibly consume assertions**: When LIGO IdPs assert identity and those assertions are used to enable LIGO users to access non-LIGO services, LIGO must trust that the SP operators will responsibly consume the identity assertion, along with any attributes asserted (with the understanding that the LIGO IdP can control precisely whether an opaque or transparent name identifier is asserted and which if any attributes are asserted on a per-relying party granularity).
- **Trust in a larger set of users**: Because federated identities make access to resources easier for users and lowers the barriers to accessing LIGO resources, LIGO can expect a significantly larger number of users to access LIGO managed SPs. LIGO must trust that each of those users will act responsibly and manage his or her federated identity appropriately. Note however that joining InCommon does not imply that all InCommon users would have access to LIGO resources, because LIGO authorization mechanisms would still be enforced. Still, increasing the number of users, whether in a federated context or not, increases LIGO's risk exposure.

Detailed Discussion of Changes to Risk Posture

Here we discuss in more detail changes to the LIGO risk posture from federating, with a specific focus on federating with InCommon, and present mitigation strategies for the risks. We group the risks into those risks from accepting external identities as asserted by non-LIGO IdPs and those of asserting LIGO identities for consumption by non-LIGO SPs:

Risks of Accepting External Identities from non-LIGO IdPs:

- **Loss of control**: LIGO's security will depend on the policies and practices of external IdPs. Any Incident response may require coordination with external IdP operators and staff.
    - *Mitigations*: Review and audit the policies and practices of external IdPs, for example the POP documents, and when appropriate for LIGO SPs requiring a certain level of assurance require use of the InCommon Silver Identity Assurance Profile (IAP). Leverage the federation community and establish contacts with the

IdP operators. Manage and centrally control with which IdPs LIGO interoperates.

- **Differing credential strength**: LIGO will not be able to unilaterally enforce a consistent credential strength and different IdPs will set different policies on password lengths and the like.
  - *Mitigations*: Review IdP practices and be selective about with which IdPs LIGO federates. Leverage the InCommon Bronze and Silver IAPs for appropriate LIGO resources.
- **Increased credential exposure**: The external identities asserted by non-LIGO IdPs will be used for many different purposes and to operate with many different resources, not just LIGO resources, so they will be subject to additional exposure from those different uses. Campus identities may be used for email, 802.1x, and a variety of web applications.
  - *Mitigations*: User training is an important mitigation strategy for all IdP operators including LIGO. LIGO should ask users and IdP operators to notify LIGO of any credential compromises (or suspected compromises). The LIGO centralized authorization infrastructure must continue to be able to block or ban specific external identities across all LIGO SPs in response to a specific incident. LIGO should develop and document its policy and procedures for a federated incident response (the authors recommend LIGO staff consider *Federated Security Incident Response Policy*[23] by the Committee on Institutional Cooperation).
- **External IdP operational security**: If a non-LIGO IdP used by LIGO collaboration members or by LIGO collaborators is compromised all LIGO users or collaborators with identities asserted by that IdP will be impacted.
  - *Mitigations*: LIGO IdM project infrastructure must implement the ability to remove trust in an external IdP quickly by updating the central LIGO metadata (accomplished by setting relatively short SP metadata refresh intervals). LIGO must monitor SP accesses for unusual behavior.
- **External IdP operational dependencies**: When a user is unable to access a LIGO resource because of an issue with an external IdP, resolution may require coordination with the external IdP operators. If the external IdP is down or otherwise unavailable, LIGO users with identities asserted by that IdP will not be able to access LIGO resources.
  - *Mitigations*: Depending on the use case (see above), some users can fall back to using LIGO identities.
- **Logging/Monitoring:** While LIGO will have full access to all logging capabilities for the SPs and the IdP it operates, LIGO will not have direct access to the logs of external IdPs, making incident response, accounting, and troubleshooting more challenging.
  - *Mitigations*: To assist in incident responses and troubleshooting efforts LIGO should configure and manage the central collection of SP logs so that helpdesk staff can more easily collaborate with their peers at other institutions.
- **Operational complexity**: Troubleshooting user problems in a federated context will be

[23] http://www.cic.net/Libraries/Technology/Federated_Security_Incident_Response.sflb.ashx

more difficult. LIGO will need the operational capacity for deciding to trust a new IdP, testing that the IdP is working correctly with LIGO SPs (including attribute release), and pushing out information about additional IdPs to SPs.

  - *Mitigations*: LIGO should continue to maintain centrally managed standard software configurations (i.e., RPMs and Debs) that LIGO SPs install and keep up-to-date. LIGO should implement SP administrator training for all web admins.

- **External identity vetting**: In the case where LIGO relies on an external IdP for asserting a name identifier it will also need to rely on the external IdP to some extent for the assertion of simple attributes such as given name, family name, and email. We expect LIGO SPs to query a LIGO attribute authority using the asserted name identifier to then retrieve assertions of “LIGO attributes” such as group memberships. To the extent LIGO relies on an external IdP for the assertion of attributes, those IdPs will have differing policies and practices for vetting those attributes. A user’s email address, phone number, and name could be vetted to different degrees (if at all) by the external IdP.
  - *Mitigations*: If LIGO requires a high LOA for an attribute then LIGO should vet it.

Risks of using LIGO identities with external SPs:

- **Password disclosure**: In the SAML protocols the user always authenticates at an IdP and so the user should never disclose a password to an SP during regular work. Users, however, can get confused and if users become used to logging in to non-LIGO services with their LIGO credentials and become complacent it may make them more susceptible to phishing attacks against LIGO passwords.
  - *Mitigations*: User training and better branding of the LIGO IdP experience so that users can become accustomed to only sending their password to the LIGO IdP.
- **Operational support**: When a LIGO user has difficulty accessing a non-LIGO SP it could (and often will) result in a support request to LIGO helpdesk staff. LIGO IdP operators will need to handle attribute release requests and other configuration details for each SP.
  - *Mitigations*: Documentation and training for both users and operational staff.
- **Information disclosure**: The disclosure of attributes and a transparent name identifier including name and email to external SPs has privacy implications.
  - *Mitigations*: The LIGO IdP’s attribute release policy controls what information is disclosed to which SPs. No information is disclosed by default or without the express approval of the LIGO policy governing the LIGO IdP. LIGO should evolve its policy on attribute release for a federated context.

Risks of Not Federating

A decision to not federate and specifically to not join the InCommon federation or any other SAML identity federation and to not federate LIGO identities also carries with it the following risks:

- **Lost or diminished scientific potential**: An arguably important fraction of non-LIGO users will decline to collaborate with LIGO if they are required to obtain and manage yet another electronic identity in order to access necessary resources. Likewise, we can expect some LIGO users to decide it is too difficult and not worth the trouble if they need to obtain and manage yet another identity in order to reach resources not managed by LIGO. Those users that do pursue yet another identity will have to become familiar with yet another mechanism for managing that identity and its credentials and how to get help and support when problems arise, slowing collaboration down and causing users more frustration.
- **Unsustainable growth and management of the LIGO Identity Management effort**: Finding qualified and high quality staff for helpdesk activities has been a significant challenge for the LIGO Identity Management project, as has been maintaining funding at the current level for such staff. Finding more funding and scaling the project to support all users with which LIGO will want to collaborate would pose a significant challenge.
- **Unresponsiveness to funding agencies**: A number of federal funding agencies, including the National Science Foundation, have signaled strong support for federated identities and the cost savings realized from organizations joining identity federations. Failure by LIGO to properly and completely analyze the benefits and risks of joining identity federations, or a unilateral declaration without strong justification of how LIGO's cybersecurity needs are different than those organizations that have already federated, might raise awkward questions at the NSF.
- **Loss of status and recognition with peers**: LIGO and the LIGO Identity Management project have to this point been considered a model example for scientific virtual organization identity management and have benefited by receiving funding from the NSF to pursue further support of federated identity within LIGO. Failure to federate and join the InCommon Federation would single out LIGO among its peers at both the NSF and in the large scientific virtual organization community as somewhat of a pariah, especially after having been recognized for its vision and leadership in the community.

Balancing and Managing Benefits and Risks

Participating in the InCommon federation and other international identity federations offers significant benefits to LIGO and the future of the gravitational wave astronomy community but at the same time increases LIGO's cybersecurity risk posture.

Trusting the IdP operators and institutions to properly manage their resources and implement appropriate credential management processes is balanced by the cost savings to LIGO of not having to scale the LIGO Identity Management project staff to manage the substantially larger user base expected during the Advanced LIGO era. Rather than hiring and training more help desk staff and system administrators so that more LIGO identities can be provisioned, LIGO can instead focus on using its existing cybersecurity and identity management staff to continually audit and measure its risk posture as part of an identity federation. More effort can be directed at implementing centralized logging and preparing for incident response and troubleshooting in a federated context rather than staffing up in order to provision and support identities for every user with which LIGO intends to collaborate.

We also note that the InCommon participants themselves have a vested interest in properly managing and securing their identity management infrastructure and processes since many participant campuses consume those same federated identities to meet their own significant demands on campus. Because of the need for many InCommon participants to provide and manage identities used to access resources that must be in compliance with FERPA and HIPPA, the level of assurance is high and can be reasonably expected to be at or beyond that which LIGO requires for access to the majority of its resources. In short, InCommon participants are motivated because many “eat their own dog food” when it comes to federated identity management.

The organizational structure, operational procedures, and existing membership in InCommon speak to the high level of professional skill being deployed to manage the InCommon SAML metadata infrastructure. Since the membership *is* the federation and the membership requires high levels of assurance to maintain the trust relationship, LIGO can be reasonably assured that the InCommon staff will be held to the highest standards. The need to trust the InCommon SAML metadata management is balanced by the work LIGO staff would have to take on in order to establish the same level of assurance and processes with each individual institution or organization throughout the world with whom LIGO would benefit from federating.

It is expected that LIGO will assert to non-LIGO SPs the “albert.einstein@LIGO.ORG” identity as a transparent name identifier along with a handful of simple attributes including first or given name, family name, common name, and email address. The risk that any non-LIGO SP will abuse these assertions, however unlikely, is balanced by the convenience to LIGO users as they use their federated LIGO identities to access important scientific resources. Further, because LIGO users will have fewer identities to manage we expect they can be trained to better manage their LIGO identity and keep it secure. The additional effort LIGO will require to troubleshoot problems and respond to incidents in a federated context is balanced by the increased collaboration opportunities for LIGO scientists.

The need to trust more users because federated identity lowers the threshold for collaboration is balanced by the LIGO mission itself since the full science potential of LIGO cannot be realized without collaboration between LIGO and other scientific communities. Since the LIGO science mission must succeed LIGO will need to trust a larger set of users one way or another. By federating with InCommon and other organizations LIGO can leverage the skills and effort of other technology staff rather than having to take on all the work itself, while at the same time using the available federation trust framework including POP documents and identity assurance profiles to continually measure and manage its risk posture.

Recommendations

Taking into account the benefits, risks, and changes to the LIGO risk posture from joining the InCommon Federation we recommend to LIGO the following course of action:

1. As a first step, join the InCommon Federation and obtain full access to the InCommon SAML metadata and the list of participant operational practices (POP) documents. Note that simply joining InCommon does not require LIGO to accept identities from any InCommon IdPs or allow use of LIGO identities with any InCommon SPs. LIGO can control federation on a per-IdP and per-SP basis.
2. Use NIST Special Publication 800-53 [24] and 800-63[25] (or equivalent as deemed appropriate by the LIGO Security Committee) as a guide and framework for classifying LIGO resources (SPs) and map the identified level of risk for each to a required assurance level.
3. Determine based on immediate collaboration needs a list of IdPs with which it would be useful to federate immediately and examine the POP documents for those IdPs and contact the IdP operators, classifying the available assurance level.
4. Based on the required assurance level for LIGO SPs, the collaboration needs, and the available assurance levels for the IdPs, develop a strategy for consuming and filtering appropriately the InCommon metadata in order to configure LIGO SPs to federate with the appropriate external IdPs. Note that we expect some high value LIGO SPs to not federate and accept external identities at all, or to require the SAML assertion to indicate that a two-factor authentication meeting particular standards has occurred.
5. Insert into the InCommon SAML metadata the metadata for the appropriate LIGO SPs as well as the LIGO IdP, with the understanding that the LIGO SPs and IdP still control locally which InCommon IdPs and SPs they will work with.
6. Develop a procedure for federated identities for vetting LIGO group memberships used for authorization to LIGO SPs.
7. Develop procedures and policies for continuing to audit the InCommon POP documents, changes in the identity assurance profile status, and the needs of the collaboration and for adding or removing federation with IdPs as necessary.
8. Develop a procedure for ongoing auditing of the use by non-LIGO SPs of asserted name identifiers and attributes by the LIGO IdP. This includes defining a process to work with any particular InCommon participant and the Federation to address concerns.
9. Develop incident response policies and procedures for federated identities, building on the *CIC Federated Security Incident Response Policy*.[26]
10. Implement infrastructure and requirements for central collection and monitoring of SP log files.

[24] http://csrc.nist.gov/publications/nistpubs/800-53-Rev3/sp800-53-rev3-final.pdf

[25] http://csrc.nist.gov/publications/nistpubs/800-63/SP800-63V1_0_2.pdf

[26]http://www.cic.net/Libraries/Technology/Federated_Security_Incident_Response.sflb.ashx

11. Develop documentation and training for helpdesk and systems staff to assist with troubleshooting issues in a federated context.

Appendix A: SAML Assertions, Protocols, Bindings, and Profiles

SAML defines XML-based assertions and protocols, bindings, and profiles. A SAML assertion contains a cryptographically signed packet of security information that a relying party (usually a SP) interprets as “assertion A was issued at time t by issuer B regarding subject S provided that conditions C are valid.” Three types of statements are provided by SAML:

1. authentication statements asserting that the principal did authenticate at a particular time using a particular method, along with other information such as the level of assurance that makes up the authentication context.
2. attribute statements asserting that the subject or principal is associated with certain attributes, which are usually transmitted as name-value pairs.
3. authorization statements asserting that the principal or subject is permitted to perform a particular action on a resource given some set of evidence.

A SAML protocol describes how the various SAML elements (including assertions) are packaged within SAML request and response elements. Generally speaking, a SAML protocol is a simple request-response protocol. An example of a SAML protocol request is an attribute query by an SP directly to an IdP over a secure back channel. SAML 2 provides the following protocols:

- assertion query and request protocol
- authentication request protocol
- artifact resolution protocol
- name identifier management protocol
- single logout protocol
- name identifier mapping protocol

A SAML binding is a mapping of a SAML protocol message onto standard messaging formats. As an example, the SAML SOAP binding specifies how a SAML message is encapsulated in a SOAP envelope, which itself is (usually) bound to an HTTP message. SAML2 specifies the following bindings:

- SAML SOAP binding
- reverse SOAP (PAOS) binding (used with the enhanced client profile or ECP)
- HTTP redirect (GET) binding
- HTTP POST binding
- HTTP artifact binding
- SAML URI binding

A SAML profile describes in detail how SAML assertions, protocols, and bindings combine to support a particular use case. Note that for SAML 2.0 most use cases involving web browser flows begin with a request at the SP and not the IdP. SAML 2.0 includes these profiles:

- Single Sign-On (SSO) Profiles
  - Web Browser SSO Profile
  - Enhanced Client or Proxy (ECP) Profile
  - Identity Provider Discovery Profile
  - Single Logout Profile
  - Name Identifier Management Profile
- Artifact Resolution Profile
- Assertion Query/Request Profile
- Name Identifier Mapping Profile
- SAML Attribute Profiles

Appendix B: Shibboleth Implementation Details

At this time the Shibboleth SP supports the following protocols or profiles:

- SSO
- attribute query
- artifact resolution
- enhanced client
- single logout
- name id management

The IdP supports the same with the exception of single logout and name id management.

Name identifiers as used by the Shibboleth SAML implementation are described by the following characteristics:

- longevity: the lifetime of most name identifiers fall in to one of three categories:
  - transient: identifiers good for a brief period of time (e.g. 5 minutes)
  - persistent: identifiers good for a long period of time (e.g. years) but which the IdP may revoke
  - permanent: identifiers good for the lifetime of an account and hence may not be revoked by the IdP
- transparency: the ability to identify a user from the name identifier. Typically the name identifier is either completely opaque or completely transparent
- targeted: whether the name identifier is intended only for a specific service provider (or group of service providers) and whether the name identifier inhibits the ability of multiple unrelated services from correlating principal activity by comparing identifier values
- revocable: whether a given name identifier can be revoked
- reassignable: whether a given name identifier, once revoked, may be re-assigned to someone else

At this time the name identifiers asserted by the LIGO IdP are permanent, transparent, not targeted, not revocable and not reassignable (we say that the name identifier is not revocable because the IdP by itself cannot and does not revoke a name identifier).

Appendix C: Shibboleth Trust Management

The Shibboleth team uses the term “trust management” when describing how Shibboleth uses cryptography to secure SAML messages and assertions not only at the level of determining whether some XML has been modified or not, but more generally to describe how Shibboleth makes the binary choice to decide if some collection of bits received is accepted, consumed, and processed or rejected. The SAML specification itself does not define or specify the cryptographic mechanisms (such as SSL/TLS or XML Signatures or similar) that could be available and used to make that binary choice. It is left to the software developers creating SAML compliant tools to incorporate and implement mechanisms meeting the necessary security needs and designing configuration controls and documentation enabling deployers and administrators to properly leverage those mechanisms and manage the cryptographic keys. Many SAML implementations rely entirely on X.509-based public key infrastructures (PKI) to implement a trust relationship and configure the tools based on that reliance. The deployment work focuses on configuring the tools to understand the trusted certificate issuers and properly consume certificate subject names. LIGO administrators have become familiar with this general approach through its common use with many grid tools including Globus Grid Security Infrastructure (GSI) and the Globus tools built on top of GSI.

The Shibboleth SAML implementation, however, takes a different approach. Shibboleth trust management is based on one or more plug-ins called “trust engines”, with each plug-in implementing a strategy for cryptographically deciding whether to “trust” a set of bits received. A feature of the plug-in mechanism is that changes to its configuration do not need to impact the rest of the system. For example, changing the trust engine plug-in mechanism does not require changes to the configuration about which attributes are sent to which relying party (assuming the relying party can interoperate with the mechanism implemented by the plug-in).

The trust engine plug-ins shipped with Shibboleth use the metadata to supply rules enforced by the trust engine to determine if a particular cryptographic key associated with some IdP or SP is trusted. At runtime the IdP or SP is configured with a set of metadata sources to use and once those sources have supplied valid metadata to the running system it is implicitly accepted and used to supply the rules to be enforced by the trust engine. In short, trust management in Shibboleth is “bootstrapped” using the metadata and so the security of the deployment depends critically on the accuracy of the metadata and its verification before being consumed by the IdP or SP.

Shibboleth supplies a trust engine plug-in that, along with properly configured metadata, enables the deployment to leverage an X.509-based PKI. It is not, however, the recommended strategy and is not the strategy LIGO has adopted.

Rather LIGO, like most other higher education institutions and research organizations, has adopted the recommendation of the Shibboleth project and leading SAML federations and uses the “inline” or “explicity key” model. The inline model is standardized at OASIS as the basis of the Metadata Interoperability Profile.

With the inline model the metadata entry for a particular IdP or SP explicitly and inline identifies the public key that the entity is authorized to use. The public key is simply detailed inside of a <KeyDescriptor> XML element. Most often (and in LIGO) an X.509 certificate is listed inside the <KeyDescriptor> element because it is a convenient vehicle for a public key but it is only the public key that is consumed and used--other parts of an X.509 certificate such as the subject or the valid dates are ignored.

This inline approach gives deployers and administrators a great deal of flexibility that help to avoid problems common to PKI. For example, the certificate in the metadata for an entity can be (and most often is) different than the certificate used by the webserver for browser-facing TLS/SSL, so that some event affecting the webserver’s X.509 certificate (such as the compromise of the CA signing certificate) need not directly impact the cryptographic trust management of the SAML assertions.

Another nice feature of the inline approach is that all of the information in the metadata with security implications is bundled and appears together--it’s easier to secure all of it in one place and then audit or review that work. In effect all of the information for a given SP or IdP in the metadata inside of the <EntityDescriptor> element acts as if it where one large and self-contained certificate and everything a relying party needs to understand to make a trust decision is encapsulated in that <EntityDescriptor>--there are no auxiliary certificate revocation lists or directory of acceptable certificate authority signing certificates as you need to have with a traditional PKI deployment.

The trade-off for using the inline model is that deployers must move the risk mitigation from the mechanisms available to a traditional PKI deployment (such as certificate revocation lists) to the metadata itself. Typically this is achieved using a “sign and expire” approach--protecting the integrity of the metadata by digitally signing it and including a validUntil attribute to limit the time for which the information can be accepted and considered valid. This approach effectively provides the same window of exposure for a compromised key that a traditional PKI using certificate revocation lists provides (the size of the window depending of course on how long the validUntil attribute or lifetime of the certificate revocation lists is configured).

Other approaches are possible. The brute force approach with metadata that is neither signed nor includes validUntil attributes relies on each administrator being contacted out of band when key replacement in the metadata is required. The “download and cache” approach (usually implemented with a cacheDuration attribute instead of a validUntil attribute) requires obtaining trusted metadata from some endpoint relatively frequently and than caching it for short periods. Note that at this time LIGO is distributing digitally signed metadata over a trusted endpoint

(HTTPS) but not yet including validUntil or cacheDuration endpoints (though the IdPs and SPs retrieve the metadata every 15 minutes).

Other trust engine plug-in models besides the "inline" model and detailed arguments about their strengths and weaknesses can be found at the Shibboleth web site.[27]

Appendix D: InCommon Structure

InCommon is a limited liability company (LLC) (not for profit) incorporated in Delaware[28] and overseen by the InCommon Steering Committee. The committee is responsible for managing the business and affairs of InCommon and its Federation, including oversight and recommendations on issues arising from the operation and management of the InCommon Federation. The committee bylaws are published online. [29]At the time of this writing the InCommon Steering Committee members include:

- Jack Suess, University of Maryland, Baltimore County – Chair
- Steve Cawley, University of Minnesota
- Joel Cooper, Carleton College – Secretary
- Mark Crase, California State University - Treasurer
- Ardoth Hassler, Georgetown University
- Chris Holmes, Baylor University
- Ken Klingenstein, Internet2 (ex officio)
- Marilyn McMillan, New York University - Vice Chair
- Kevin Morooney, Penn State
- John O'Keefe, Lafayette College
- Craig Stewart, Indiana University
- Shel Waggener, University of California Berkeley

Appendix E: Overlap between LSC and InCommon

The complete list of InCommon participants is available online.[30] At this time the following institutions are InCommon participants and have an active group in the LIGO Scientific Collaboration:

- California Institute of Technology
- California State University, Fullerton
- Carleton College
- Columbia University
- Louisiana State University
- Massachusetts Institute of Technology

[27] https://wiki.shibboleth.net/confluence/display/SHIB2/TrustManagement

[28] http://www.incommonfederation.org/docs/policies/InCommonLLC.html

[29] http://www.incommonfederation.org/docs/policies/InC_SCbylaws.html

[30] http://www.incommon.org/participants/

- Northwestern University
- Penn State
- Sonoma State University
- Stanford University
- University of Florida
- University of Maryland
- University of Massachusetts Amherst
- University of Michigan
- University of Minnesota
- University of Oregon
- University of Texas at Brownsville
- University of Texas at Austin
- University of Wisconsin-Milwaukee

The following InCommon participants are of interest because of the likelihood of collaboration between members of these institutions and LIGO:

- Argonne National Laboratory
- Cornell University
- ESnet
- Fermilab
- Georgia Institute of Technology
- Internet2
- National Science Foundation
- Ohio State University
- Pacific Northwest National Laboratory
- Princeton University
- University of Illinois at Urbana-Champaign
- University of Southern California
- University of Wisconsin-Madison

Appendix F: IAP Identity Assurance Criteria

The table below, taken directly from the IAP document, summarizes all of the identity assurance criteria defined for Bronze and Silver IAPs.

Functional area	Criteria	Bronze	Silver
Business, Policy and Operational criteria	InCommon Participant	✔	✔
	Notification to InCommon	✔	✔
	Continuing Compliance	✔	✔

Registration and Identity Proofing	RA authentication	n/a	✔
	Identity verification process	n/a	✔
	Registration records	n/a	✔
	Identity proofing	n/a	✔
	Existing relationship	n/a	✔
	In-person proofing	n/a	✔
	Remote proofing	n/a	✔
	Address of Record confirmation	n/a	✔
Credential Technology	Credential unique identifier	✔	✔
	Resistance to guessing Authentication Secret	✔	n/a
	Strong resistance to guessing Authentication Secret	n/a	✔
	Stored Authentication Secrets	✔	✔
	Protected Authentication Secrets	✔	✔
Credential Issuance and Management	Credential issuance process	n/a	✔
	Credential revocation or expiration	n/a	✔
	Credential renewal or re-issuance	n/a	✔
	Retention of Credential issuance records	n/a	✔
Authentication Process	Resist replay attack	✔	✔
	Resist eavesdropper attack	✔	✔
	Secure communication	✔	✔
	Proof of Possession	✔	✔
	Session authentication	✔	✔
	Mitigate risk of sharing credentials	✔	✔
Identity Information	Identity record qualification	✔	✔

Management			
Assertion Content	Identity Attributes	✔	✔
	Identity Assertion Qualifier	✔	✔
	Crytographic security	✔	✔
Technical Environment	Software maintenance	n/a	✔
	Network security	n/a	✔
	Physical security	n/a	✔
	Reliable operations	n/a	✔