A Roadmap for using NSF Cyberinfrastructure with InCommon

A practical guide for using InCommon and Identity Federation to support NSF Science and Engineering

William Barnett, Craig Stewart, Alan Walsh, Von Welch

Indiana University

Please cite as:

Barnett, W., Stewart, C.A., Walsh, A., and Welch, V. A Roadmap for Using NSF Cyberinfrastructure with InCommon. DOI: ###___. Available from: http://hdl.handle.net/###___ and http://www.incommon.org/nsfroadmap.html

About This Document

This document provides a Roadmap for using the InCommon identity federation to enable researchers to access NSF cyberinfrastructure (CI) via their campus authentication service. It presents benefits and challenges of using InCommon for NSF cyberinfrastructure, and guidance in overcoming the challenges. The Roadmap has three main sections, each aligned for a different audience:

A. *Benefits, Challenges and Overview* is intended for campus and project leadership, scientists and engineers using CI. It provides a summary of InCommon, relevant technologies and the benefits and challenges their adoption brings.

B. The *Guide to Technical Deployment* is intended for information technology professionals, from campuses and NSF cyberinfrastructure projects, and is a guide for deployment of InCommon software and services.

C. The *Guide to Policy and Business Processes* is intended for managers and policy makers, and discusses the policy, privacy, financial and other factors of InCommon deployment. Again it is both for staff from campuses and NSF cyberinfrastructure projects.

A final section provides a glossary, references and other resources.

In order to be insulated from inevitable changes in technologies and to be as comprehensible as possible, the document avoids capturing technical details when it can, instead providing references to existing (particularly online) documentation provided by InCommon, Internet2 and other organizations.

Document Scope

There are a wide variety of federated identity technologies and organizations that seek to form trust amongst organizations for online collaboration. This document is specific to InCommon, with its focus on higher education and research institutions, institutions that are highly aligned with the NSF science and engineering community.

This document also focuses on the needs of NSF cyberinfrastructure (CI) Projects, which are projects providing computer-based resources (e.g., compute cycles, data resources, shared instrumentation, web-based applications, virtual organizations) to scientists and engineers, and having some need to identify those researchers in order to, for example, perform access control, resource authorization, audit usage, or provide personalization. A full discussion of CI is beyond the scope of this document, for context the reader is referred to [59]. As subsequently discussed in Section A.1, NSF CI projects frequently have requirements above and beyond normal InCommon service providers and this document focuses on meeting those requirements.

In addition, the document is scoped as follows:

- InCommon is most accurately a federation based on the SAML protocol, and this document has chosen to focus on Shibboleth as a popular open source SAML implementation used in InCommon. Alternatives to Shibboleth, InCommon and SAML are discussed in Section A.5.
- As discussed in the Guide to Policy and Business Processes, InCommon allows for higher levels of assurance beyond the base level required for membership – i.e. Bronze and Silver. For the purposes of brevity, this document constrains itself to a brief discussion of when these higher assurance levels may be appropriate for a CI project to consider.
- This document covers cyberinfrastructure projects serving NSF researchers and institutions of higher education and research that host those researchers. Effort was made to discuss experiences with a variety of institutions of different sizes as to avoid assumptions regarding available resources and expertise.

Acknowledgements

The authors thank the following individuals who volunteered their time to serve as the editorial board for the development of this document and provided invaluable feedback and suggestions: James Basney (NCSA/U. of Illinois), Michael Beyerlein (Purdue U.), Ken Klingenstein (Internet2), and Michael McLennan (Purdue U.). Ken Klingenstein also contributed much of the text for the Jean Blue use case in the Roadmap.

The authors thank the NSF Advisory Committee on Cyberinfrastructure Campus Bridging Task Force for their guidance on this document.

The authors thank the following individuals for sharing experiences and suggestions, which greatly improved this document: David Banz, Matt Kolb, Redmond Militante, and John O'Keefe.

The authors thank the InCommon staff for their support and the excellent materials they are producing, which were invaluable in authoring this document. Additionally, Tom Scavo of InCommon provided valuable feedback on several occasions during the writing of the document.

This material is based upon work supported by the National Science Foundation under OCI Grant No. OCI-1040777. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Table of Contents

A WHY USE INCOMMON AND FEDERATED IDENTITY 9

A.1 WHAT IS UNIQUE ABOUT NSF CI? 10
A.2 BRIEF OVERVIEW OF FEDERATED IDENTITY AND INCOMMON 11
A.3 BENEFITS FOR RESEARCHER, INSTITUTION AND CI PROJECT 13
A.4 CHALLENGES OF FEDERATED IDENTITY 17
A.5 ALTERNATIVES TO INCOMMON AND SHIBBOLETH 20
A.6 SECTION CONCLUSION 21

B GUIDE TO TECHNICAL DEPLOYMENT 23

B.1 INTRODUCTION TO TECHNICAL ISSUES 24
B.2 TECHNICAL DEPLOYMENT FOR INSTITUTIONS (IDENTITY PROVIDERS) 27
B.3 TECHNICAL DEPLOYMENT FOR CYBERINFRASTRUCTURE PROJECTS (SERVICE PROVIDERS) 30

C GUIDE TO POLICY AND BUSINESS PROCESSES FOR DEPLOYMENT 42

C.1 INTRODUCTION TO POLICY AND BUSINESS PROCESS ISSUES 42
C.2 EFFORT REQUIRED FOR SHIBBOLETH DEPLOYMENT AND INCOMMON MEMBERSHIP 45
C.3 INSTITUTIONAL DEPLOYMENT: POLICY AND BUSINESS PROCESS ISSUES 49
C.4 CYBERINFRASTRUCTURE DEPLOYMENT: POLICY AND OTHER ISSUES 53

D GLOSSARY OF TERMS 57

E REFERENCES 61

F ADDITIONAL RESOURCES 65

F.1 FUTURE RESOURCES 65
F.2 IDENTITY MANAGEMENT RESOURCES 65
F.3 RESOURCES FOR FEDERATED IDENTITY DEPLOYMENT 66

A Roadmap for using NSF Cyberinfrastructure with InCommon

Benefits, Challenges, and Overview

Abstract

Benefits, Challenges and Overview is intended for campus and project leadership, and scientists and engineers using cyberinfrastructure. It provides a summary of InCommon, relevant technologies and the benefit their adoption brings to campuses supporting researchers, the researchers themselves, and cyberinfrastructure deployments.

A Why Use InCommon and Federated Identity

> "Today's scientists and engineers need access to new information technology capabilities, such as distributed wired and wireless observing network complexes, and sophisticated simulation tools that permit exploration of phenomena that can never be observed or replicated by experiment. Computation offers new models of behavior and modes of scientific discovery that greatly extend the limited range of models that can be produced with mathematics alone, for example, chaotic behavior. Fewer and fewer researchers working at the frontiers of knowledge can carry out their work without cyberinfrastructure of one form or another."

As this quote from the National Science Foundation's (NSF) "Cyberinfrastructure Vision for 21st Century Discovery" [59] describes, cyberinfrastructure (CI) is a key and necessary component to support increasingly collaborative science and engineering. As opposed to traditional high-performance computing, a key goal of CI is to support scientific collaboration through a variety of computational, network, data and software elements distributed across campuses, regional, national and international organizations, and spanning scientific communities.

Critical to supporting the CI ecology is a well-coordinated, usable identity management system on which CI services can be built to allow for trusted collaboration and sharing of compute and data resources across researchers' institutions. To this end, the joint EDUCAUSE-CASC workshop on CI [13] recommended:

> "Agencies, campuses, and national and state organizations should adopt a single, open, standards-based system for identity management, authentication, and authorization, thus improving the usability and interoperability of CI resources throughout the nation."

The same workshop report continues and specifically recommends the InCommon federation as the current best solution for broad adoption.

The InCommon federation represents an implementation of *federated identity*. Federated identity refers to the practice of one organization receiving and utilizing identity information regarding a user from another organization, typically the organization at which the user is employed or is otherwise a member. The objective is that the latter organization leverages the work the first organization has done in enrolling the user, managing a credential (e.g., password[1]) for the user, and asserting attributes about the user.

[1] We note that campuses are free to use any authentication credential they desire with InCommon, however passwords are common and this document tends to use that term, as it is familiar for many readers.

[2] The goal of uApprove [93] is to change attribute release from an institutional policy to a decision by

Federated identities in general, and InCommon in particular, are becoming standards in establishing trust in the research sector. InCommon has other federal partners, including the Department of Energy's Energy Sciences Network (ESNet) and the National Institutes of Health.

The goal of this Roadmap is to encourage more effective scientific collaboration and team science supported by campus and NSF CI by fostering the use of InCommon in order to:

1. Allow researchers to more easily collaborate and coordinate multiple resources through a single identity system rather than spending effort on managing multiple identities.
2. Allow NSF CI projects to leverage InCommon saving effort spent on establishing their own identity systems.
3. Allow campuses and other institutions to provide their researchers with a consistent identity system for local research and administrative computing, and remote research computing.

The Roadmap strives to achieve this goal by providing campuses and CI projects with the rationale and guidance for deploying and using federated identity, joining InCommon, and supporting collaborative science using that infrastructure.

A.1 What is unique about NSF CI?

A reasonable question is why NSF CI needs a roadmap in addition to the guides for adoption of federated identity and InCommon that already exist? NSF CI represents a number of science-enabling collaborations and resources, including rare (even unique) and valuable computational, data and instruments. CI representing these resources often has one or more of the following attributes, which make them atypical of InCommon service providers:

- Strong requirements for secured sharing: Computational resources are commonly among the worlds most powerful and it is not unheard of for them to fall under U.S. Export Control law. NSF CI also manages scientific data created and owned by researchers, data which can have privacy, integrity and trusted sharing requirements based on its implications to research results that can effect scientific standing and policy issues (e.g., climate change, human subjects information).
- Distributed researcher communities: A NSF CI project typically has distributed, dynamic researcher communities that don't conform to any group of researchers at any one campus or other institution. For example, access to TeraGrid is granted via a national allocations process that occurs multiple times per year [66]. Many projects have less formal processes involving collaboration participants who may come and go depending on current research interests and their alignment with the project.
- A history of identity management: Because of the nature of their resources and communities, NSF CI projects often have stringent, self-managed access

control requirements. To meet these requirements, there is a history in NSF CI projects of performing strong vetting of their users and persistent account management. This creates a situation of researchers having multiple digital personas (one for their institution plus additional personas for each project they are involved in), thus creating a barrier to trusted virtual collaboration.

- A need for incident response: NSF CI projects often have a need to perform incident response to understand the implications of any data breech; a need that is otherwise underrepresented in typical federated identity applications.
- Non-web access modalities: NSF CI projects often have command-line access modalities that are not currently supported by typical federated identity software (though as we discuss in Section F.2, such support is planned). For example, a common means of accessing NSF CI is through secure shell (SSH) to obtain command-line access and do job submission.

A.2 Brief Overview of Federated Identity and InCommon

We briefly present some basic terminology regarding federated identity and InCommon as shown in Figure 1. For more complete and technical definitions of the terms, the reader is referred to the Glossary.

The term "**federated identity**" refers to the ability to utilize a user's identity, as managed by one organization, across multiple organizations. A collection of organizations that agree to a common set of practices and policies for federated identity are referred to as a **federation**, with the member organizations being referred to as **participants**.

An example of a federation is **InCommon**, which focuses on institutions of higher education and organizations providing services to those institutions. InCommon is governed by its members [25] and operated by Internet2.

Within a federation, participants are **identity providers** that instantiate institutionally managed services that authenticate users and allow their identities to be shared with **service providers**, who consume those identities in order to provide access to resources or services. For example, the Indiana University identity management system represents an identity provider, providing institutional credentials and guaranteeing that researchers with Indiana University logins have been physically vetted. A service provider, such as the Indiana Clinical and Translational Sciences Institute HUB [43], accepts institutional credentials from a number of identity providers and allows users of those identity providers access to cyberinfrastructure services such as data management and shared computational facilities.

The term "**identity**" is used to refer the aggregate of **identifiers**, which uniquely identifies an individual, with a collection of zero or more **attributes** regarding that person. Identifiers can be ephemeral, used only for a single session, pseudonomymous, persistent for arbitrarily long periods of time but not reflecting the user's physical identity, or fully identifying, persistent and reflective of the user's physical identity (e.g., an email address). Attributes provide information about a person such as their institutional role (e.g., faculty), department, class enrollment, or contact information (e.g., phone number). **Privacy** is preserved by the controlled release of identity information to service providers, a process referred to as **attribute release**.

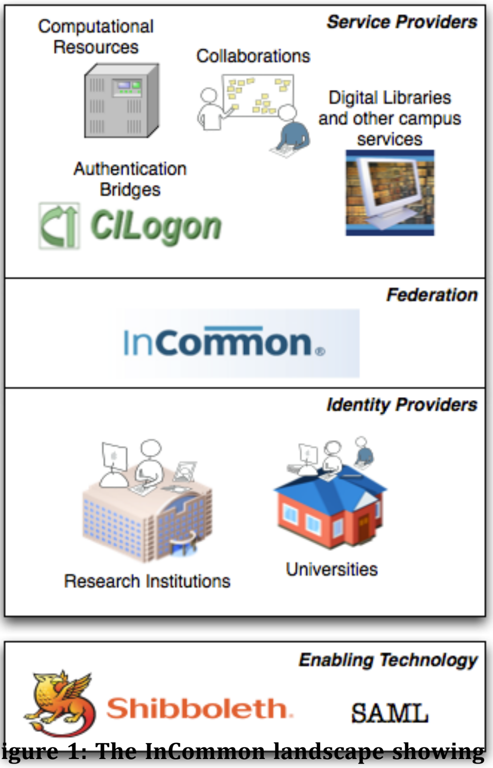

Figure 1: The InCommon landscape showing Identity Providers (campuses and institutions), the InCommon Federation, and Service Providers such as digital libraries, campus services, collaboration, and cyberinfrastructure. Enabling technologies include the SAML standard and the Shibboleth software.

InCommon is based on the **SAML** standard [67], which defines message formats and protocols to provide for interoperability among participants. Building on SAML, **eduPerson** [15] defines a set of user attributes common to educational institutions that is heavily used in InCommon.

A key function of the federation is to manage and distribute **metadata** among its participants. Metadata, whose format is defined by the SAML standard, is information that describes federation participants (identity and service providers) and allows participants to securely communicate identity information.

To utilize InCommon, software is needed that implements the SAML standards and provides identity providers with the tools to provide identities, service providers with the tools to consume identities, and users of the system the tools to express their intents with regards to authentication and privacy. A number of commercial and open-source SAML implementations are available. **Shibboleth** [78] is frequently used in InCommon. It is freely available as an open source project spearheaded by Internet2, and the focus of choice for this Roadmap.

A.3 Benefits for Researcher, Institution and CI Project

In this section, we describe the benefits for using federated identity and InCommon to support NSF science and engineering from three perspectives: that of the NSF researcher, that of the CI project, and that of the researcher's institution.

A.3.1 Benefits to the Researcher

To help understand the benefits of federated identities in research, we introduce you to Jean Blue, Professor and Researcher, and present a morning in her life supported by federated identity.

> Dr. Blue gets up in the morning and logs into her campus to check her email. One of the notes is from her campus sponsored research office, indicating that a report is due on her NSF grant. She goes to the sponsored research office web site, and selects the research.gov link there. Because she previously logged in to her campus to check email, and because research.gov trusts her campus to provide accurate, up to date identity information, Dr. Blue's prior authentication is automatically used to allow access to her research.gov account and Dr. Blue uploads the requested report.
>
> Another one of her emails alerts her to new data posted on the translational research wiki at National Institutes Health (www.ctsawiki.org/). She navigates to wiki, which like research.gov uses her prior institutional login to authenticate and welcome her directly to her personal wiki page. Seeing new data sets available, she decides to launch a job on the TeraGrid to analyze them. She opens a browser window to CILogon (cilogon.org), which notes her campus authentication but asks her to release some additional attributes, such as a screen name, as requested by the CI service providers.
>
> Jean then checks on the latest data for a clinical trial she is managing. The data is stored on Jean's local campus and accessible via secured web site, which permits her access based on her previous login. The site presents her with a request for access from a colleague at another institution to collaborate on a paper they are co-authoring. To make the request, the colleague authenticated to that data store with their campus login and approved the release of attributes - campus department and role in this case - to help validate the request. Jean reviews the request, recognizing the collaborator based on their name and attributes, and approves the request, granting access without having to create another username and password for the colleague.
>
> Finally, Jean jumps over to Elsevier (www.sciencedirect.com) to check some recent journals. The site welcomes her back, granting her access based on her status as her campus without knowing her actual identity, and alerts her that three of her watch-list words had been triggered by

> articles in her chosen journals. Jean sighs, and flags them for later reading.
>
> It has been a busy morning, with a lot of collaboration done, all with a single campus identity.

How much of Jean Blue's story is real today? Every site with a URL is operational today using federated identities; the other scenarios are under active development.

As illustrated by this example, the direct benefit to the researcher is that they can utilize many CI resources without having to create yet another username and password for each. Initially this expedites obtaining access to CI by removing delays with secure distribution of passwords to these resources. Over the lifetime of the researcher's access, it removes the need for the researcher to manage a separate username and password, reducing the chance of forgetting the password and giving them an existing campus support system for changing the password, resetting it in the event they forget it, etc. This not only means that there is a higher level of security, but also less overall effort since each of these services does not have to repeat a vetting process to ensure that the researcher is who they claim, instead leveraging the effort performed by their institutional identity provider. This is especially important for access to secured resources such as the TeraGrid or sensitive data, such as human subjects data.

In the bigger picture, the utilization of their campus login for access is a key first step to allowing someone to utilize any CI without concern about where it might be located or who is operating it. This allows researchers to focus on science and scientific collaboration without having to worry about what collaborators have accounts where, setting up authentication services, and the like.

For researchers with security concerns about data and other resources they are sharing in their collaboration, the use of campus credentials provides greater assurance, as collaborators will be less inclined to share or otherwise mishandle those credentials as they might a password generated solely for the collaboration. The credentials are also tied to the collaborator's position at an institution, meaning that in the event a researcher loses academic status, and the identity will be revoked and cannot be used for access. This allows service providers to more easily provide trusted access to sensitive data, and administrative processes for study review, like Institutional Review Boards (IRBs) can be undertaken with greater confidence and streamlined.

Finally, funding agencies, such as NIH (see [49]) and NSF, have joined InCommon and are moving towards federated identity as the access mechanism for grant application and administration. Utilization of federated identity for CI will bring uniformity to the authentication mechanism for science in line with the business processes of doing science.

A.3.2 Benefits for the CI project

> “Harvesting the science content from LIGO [Laser Interferometer Gravitational-Wave Observatory] data is a collaborative effort between instrumentalists, data analysts, modelers, and theorists. Efficient collaboration begins with scalable and robust identity management infrastructure that can easily be leveraged and integrated with the wide spectrum of tools LIGO scientists use to collaborate and analyze the LIGO data. Middleware from Internet2, including Shibboleth and Grouper, is enabling more LIGO science through easier collaboration and access to resources.” -- Scott Koranda, Senior Scientist at the University of Wisconsin-Milwaukee and lead architect of the LIGO [54] Identity Management effort

A NSF CI project receives many of the same benefits from InCommon as any other InCommon Service Provider. Descriptions of these benefits, including multi-media presentations, can be found at the InCommon for Service Providers web site [32]. We summarize the benefits here and highlight those most applicable to CI projects.

The immediate benefit of federated identity to a project with any sort of access control requirements is that they still control who has access to their resources, but authentication is performed by their researchers’ home institutions, getting the project out of the business of creating password databases and distributing passwords (and re-distributing them when they are lost). Initially, this has the benefit of expediting the granting of access to new users since they already possess their passwords. A case study from the Swedish Alliance for Middleware Infrastructure on federated identity addressing costs of the identity vetting process can be found in [55].

In the longer term, federated identity also reduces overhead on the project for managing researchers’ passwords – e.g., resetting forgotten passwords, regular expiration – allowing the researcher instead to use already familiar campus processes. This reduction in responsibility can be of particular benefit to smaller, resource-constrained projects and collaborations.

From a security perspective, the use of the campus password for authentication also decreases the chance the researcher shares or otherwise mishandles that password, resulting in increased assurance of the user’s identity. Removing the need to distribute passwords reduces risk of password exposure. And expediting researcher access by removing the need for password distribution acts to decrease the motivation for users to share passwords.

Furthermore, access can be based on researcher’s attributes; for example, their role as faculty at their campus, either solely or in addition to the user’s identifier. This use allows for automatic provisioning and de-provisioning of researcher access without time consuming verification of these attributes by project staff. For example, a service could verify on every use that a researcher remains their position as asserted by their home institution.

From the perspective of adoption, providing researchers access with an existing credential, and one potentially in use by other CI projects, removes one step in setting up the project CI, reducing a barrier to entry and encouraging use.

A.3.3 Benefits for the researcher's institution

As in the previous section on benefits to CI projects, campuses receive a number of benefits from the adoption of InCommon and federated identity that are documented by InCommon [28]. We summarize those benefits here and highlight those most applicable to supporting NSF science and engineering CI projects:

- ***Controlled, scalable access to external services.*** Shibboleth and InCommon provide a scalable means of providing controlled access to external services. For example, they can replace current schemes based on IP addresses for controlled access to digital libraries with a scheme based on the institution's provisioned user base [37]. A complete list of InCommon Sponsored Partners either providing or in the process of providing access via InCommon can be found on the InCommon participants web page [11].
- ***Privacy controls.*** Shibboleth gives the campus and its faculty, staff and students privacy controls with regards to what attributes are released to each service provider. It supports anonymous and pseudonymous authentication, and the ability to receive user consent for the release of attributes, which can be beneficial in addressing legal requirements such as FERPA or HIPAA.
- ***Visibility into CI usage.*** The use of federated identity gives the campus visibility into the use of CI (and other services) by its user community since the campus is now part of the authentication process. This allows for the collection of aggregated, privacy-respecting statistics on what services are used by what types of users, and with what frequency.
- **Grant competitiveness.** Supporting federated identity will increasingly be important to grant competitiveness as the grant process moves to InCommon, as science increasingly moves to team science, and as effective collaborations improve science outcomes. InCommon will permit institutional researchers improved, or even preapproved, access to offsite data and analytical resources, allowing them to be more competitive in terms of research.
- ***Uniform authentication mechanism.*** Providing an authentication mechanism usable by both researchers on campus and their external collaborators helps prevent "home-grown" authentication systems being set up by researchers in front of potentially sensitive data (e.g., a collaboration sharing clinical data). In general, providing the same authentication mechanism for internal CI that is used by external CI allows the campus to provide CI locally for researchers and their collaborators that removes a barrier to transitioning between that local CI and regional or national CI.
- ***Internal single sign-on.*** Federated identity provides web single sign-on internal to the campus with the usual benefits of doing so, namely a single password for users, centralized provisioning of accounts, and central auditing.

- ***InCommon certificate service.*** A side benefit to joining InCommon is access to the InCommon Certificate Service [29], providing X.509 certificates (SSL, EV, personal signing, encryption, and code signing) for a fixed annual fee.

A.4 Challenges of Federated Identity

In order to be balanced in our presentation, we discuss here the challenges to deploying and using federated identity and InCommon. In the following section, we discuss some of the alternatives to InCommon and their trade-offs. The authors of this Roadmap believe these challenges are out-weighed by the advantages and the approach of this roadmap is at least as good a choice as the alternatives, but we acknowledge that every solution has disadvantages as well as advantages and so include this section in the interest of full disclosure.

A.4.1 Mature Identity Management as a Required Prerequisite

In our discussions with organizations that have deployed Shibboleth and joined InCommon, a consistent prerequisite that came up was the organization having a "mature" identity management system in place before it undertakes federated identity. What constitutes "mature" is somewhat subjective, however the following have emerged as key features:

- *A centralized user directory infrastructure*. The organization has a single known, authoritative source for user information (authentication and attributes) with defined interfaces for accessing that information and controls on its modification.
- *Understood business processes for user enrollment*. The organization understands how users are enrolled in their identity management system, how their roles are assigned, and how they are removed from the system. This includes an understanding, at least, of what the edge cases are; for example: guest logins, anonymous library users, contractors, incoming students, and incoming faculty.
- *Automated user provisioning*. Based on the business processes, user provisioning and de-provisioning in the identity management system (i.e. addition, removal and attribute management of users), should be, at least for a majority of users, automated.

To be clear, an organization doesn't need to have these completely solved (no organization probably does), but more complete solutions lead to easier federated identity deployment and higher levels of trust.

Establishing an identity management system is outside the scope of this document, however some resources for doing so can be found in Section F.2.

A.4.2 Changes to Risk Profile

Federated identity turns what used to be an identity management process that was internal to an organization into a process distributed across multiple organizations. This brings changes to the risk profile of an adopting organization:

- *Reliance on the external infrastructure*. For a CI project, the trade-off for reduced workload and interoperability is a reliance on the InCommon federation and federation partners (and interconnecting infrastructure), which entails risks to both reliability and security. Related to this is that in the bigger picture, by increasing the scope of use for a single authentication, we increase the impact if that authentication is fraudulent (put simply, if the researcher's campus password is stolen, it grants illicit access to more services with federated identity). Quantification of these risks is difficult because they depend on the specific set of services used by each individual researcher and a lack of long-term operational data, but is something participants need to be aware of and accept (or identify mitigation strategies for).
- *Reliance on enabling technologies*. The use of federated identity involves relying on enabling technologies, for example Shibboleth software. Mitigating this risk is InCommon's use of open standards and Shibboleth's track record as an Internet2 member-supported software project.
- *Risk of user attribute exposure*. Shibboleth provides attribute release policies to control, on a service provider by service provider basis, the sharing of user attributes. Nevertheless, there is still a risk of human or software error resulting in inappropriate sharing. Emerging technologies such as uApprove [93] allows users to participate in attribute release and mitigates this risk.

A.4.3 Expenses of InCommon Membership and Shibboleth Deployment

For organizations that chose to deploy Shibboleth and manage the process of joining InCommon themselves, which is a very typical thing to do, the largest cost will be staff time. In the subsequent section (A.4.4) we summarize the effort required for organizations to estimate this cost.

In addition to staff time other expenses include:

- InCommon Participant Fees: Currently $1000-$3000 annually depending on the size of the organization plus a $700 one-time fee. Please see the InCommon web site [31] for details and changes since the writing of this document.
- Web certificates for identity and service providers. As with any other secure web server, these services need web server certificates. (Note that organizations could use the InCommon Cert Service as described in Section A.3.3 for these certificates.)

Alternatively, organizations can choose, as discussed in Section A.5, to outsource portions of the Shibboleth deployment - from design consultation to service hosting. This obviously shifts internal effort re-allocation to out-of-pocket expenses, and while organizations may choose this route, it does not appear to be a requirement for most organizations capable of running their own identity management systems. Outsourcing identity management services can also create additional risks, such as an outside entity having possession of institutional credential information.

A.4.4 Effort Required for InCommon Membership and Shibboleth Deployment

Most organizations choose to deploy Shibboleth (or an alternative) and manage joining InCommon themselves. As discussed in the previous section on expenses, staff time is the largest expense of this approach. It is difficult to give a quantified effort level for participating in federated identity as processes, expertise, culture and other factors vary between organizations and projects. We instead break down in Table 1 the effort required for deploying and maintaining federated identity and InCommon membership into a set of equivalencies to other common activities in terms of required effort and skills. The expectation is that the reader can judge the effort that these equivalent activities would require for their organization or project, and translate that into a quantified estimate for participation in InCommon.

Note that we provide only a summary of the tasks in this section, focusing on the effort level rather than "how to" details; for details on accomplishing the tasks, please see the subsequent Roadmap sections on Technical Issues, and Policy and Business Process Issues.

InCommon Membership Activity	Roughly Equivalent Activity/Effort
Leadership for process of joining	Requires CIO or delegate with support of campus leadership.
Policy and business process documentation and modification	Major authentication policy change, e.g., establishing a new minimum password strength.
Signing InCommon membership agreement	Contract signing.
Deployment of Shibboleth Identity Provider software	Deployment of a web single sign-on system (e.g., CAS [5])
Deployment of Shibboleth Service Provider software	Deployment of a web application protected by web single sign-on; varies greatly by application.
Addition of a federated partner	Technically is a minor configuration change. From a policy perspective varies based on partner's requirements; having well defined process in place eases this.
Software/service maintenance	Maintaining a web single sign-on service. A few additional activities are minor overhead.

Table 1: Activities involved in joining and maintain membership in InCommon and rough estimates of the effort required based on equivalent activities.

A.5 Alternatives to InCommon and Shibboleth

We briefly describe some alternatives to the InCommon and Shibboleth approach highlighted in this Roadmap, and discuss their trade-offs.

- *Bilateral agreements without InCommon*. It is possible, at least in theory, to forgo a federation and use a set of bilateral agreements to support a federated identity fabric. Given the relatively low cost of supporting InCommon, the time costs of establishing similar bilateral agreement would seem to quickly outpace any savings.
- *Using social networking identities*. Instead of InCommon, an organization or project could utilize identities as asserted by social networking sites (e.g., Facebook, Google, Yahoo) using technologies such as OAuth [68] and OpenID [96]. The advantages and disadvantages of this approach is an area of some debate currently. On the side of social networking is that social networking sites absorb the costs of providing identities and users tend to already have such accounts. On the other hand, social networking identities tend to be self-asserted by the users (e.g., see [17]). There is no institutional authority behind them, thus InCommon has the potential for higher strength of authentication. InCommon has the advantage of greater stability provided by higher education institutions, as opposed to commercial entities, which may change their practices due to business concerns. InCommon also has the ability to include attributes from the user's home institution. It is also not an either-or situation, use cases are emerging [50] where these technologies are complementary: Shibboleth is used to provide stronger authentication for employees and students, and OpenID is used for guest accounts to access less-sensitive resources.
- *Projects can establish their own identity management system*. CI projects can establish their own identity management systems, even utilizing single sign-on solutions to achieve some benefits of federated identity (such as the Earth Systems Grid [88] has done). This approach brings the benefit of being more of a known approach and keeps the project in control of their destiny, at the cost operating their own authentication infrastructure and a lack of interoperability.
- *Alternative SAML implementations*. There exist a number of open source and proprietary implementation alternatives to Shibboleth. We do not try to capture a list of such implementations here due to the fact it would be quickly out of date, but the list of InCommon affiliates [26] would be a good starting point for researching these alternatives. Organizations may want to explore these options, as it is certainly possible that while Shibboleth serves many organizations well, an alternative may serve a particular organization better. For example, an organization heavily using Microsoft products should explore federated identity products offered by Microsoft.
- *Utilize a third-party identity provider*. There exist commercial parties that can provide federated identity provider services that interoperate with

InCommon for an organization that does not want to deploy their own service. Based on discussions, we believe a decision to pursue such an option is based more on an organization's culture than any technical or effort consideration. The list of InCommon affiliates [26] and sponsored partners [11] are good places to start exploring options.

A.6 Section Conclusion

This concludes the first section of the Roadmap for using NSF Cyberinfrastructure with InCommon. We hope that it has provided a good overview of InCommon, federated identity, and the advantages, disadvantages and challenges of deploying a federated identity system to support collaborative research and enable better science outcomes. This document has two subsequent sections: one on Technical matters and one on Policy and Business Processes that go into more depth on addressing the challenges involved in joining InCommon and using it to support NSF cyberinfrastructure.

Two versions of this Roadmap are distributed: A complete version and, mainly intended for print, an abbreviated version. The abbreviated version does not include the two subsequent sections. They be may found online at:

http://www.incommon.org/nsfroadmap.html

A Roadmap for using NSF Cyberinfrastructure with InCommon

Guide to Technical Deployment

Abstract

The *Guide to Technical Deployment* is intended for information technology professionals, from campuses and NSF cyberinfrastructure projects, and is a guide for deployment of InCommon software and services.

B Guide to Technical Deployment

Part of implementing federated identity is the deployment and operation of technical services that handle the transmission of identity information from the researcher's institution to the project or resource that utilizes that information. The goal of this section is to provide direction for the deployment and operation of these services for both the researcher's institution and the CI project, along with their integration with the existing services at those organizations to enable their use.

This section is split into guidance for the researcher's institution (the identity provider) and for the CI project (the service provider). Since Shibboleth deployment and joining InCommon are well documented by the Shibboleth project and InCommon respectively, this roadmap covers the generic aspects of doing so briefly and focuses on aspects to support NSF CI.

Details specific to supporting NSF CI are highlighted, as this paragraph is, to allow users familiar with Shibboleth and InCommon to quickly skim and locate these steps.

Note that a typical deployment process, for both an identity provider and a service provider, is to go through the deployment process once to deploy a prototype service to be tested by a small number of friendly users and staff, digest the lessons learned from that experience, and then plan out a production deployment. We recommend that approach, as difficulties with Shibboleth deployments tend to lie in its interactions with other services. This approach will expose those problems as early as possible in the deployment process.

B.1 Introduction to Technical Issues

We briefly introduce the technical issues in this section that span both identity and service providers.

B.1.1 Attribute Release and Persistent User Identifiers

A strength of Shibboleth is its ability to release attributes in a controlled manner from identity providers to service providers. When a participant joins InCommon as a service provider, they undergo what is often referred to the "boarding process" [46]. This process entails that service providers determine their attribute needs, request those attributes of the identity providers representing their users, and then the identity provider administrator configures what attributes will be released to the service provider. For background on attribute release, see [52]. This process has both policy and technical aspects; in practice, the effort required for the policy aspects, which we discuss in Section C on Policy and Business Practices, eclipse the effort required for the technical aspects discussed in this section.

In practice, the attribute of interest to NSF CI that is most unusual, though not unique, is a persistent user identifier so that identity-based access control and auditing can be implemented.

Within InCommon, with its use of the eduPerson attributes, there are two typical ways of accomplishing the release of a persistent identity:

- Use of the eduPerson Principal Name (ePPN). In this scenario an internal identifier for a user is used to generate an identifying attribute that looks very much like an email address (and could actually be an email address). Directions for configuring ePPN in the context of Shibboleth can be found at [77].
- Use of the eduPerson Targeted Identifier (ePTID). In this scenario a unique identifier is generated for the user for each relying party they visit. Directions for configuring ePTID in the context of Shibboleth can be found at [77].

A possible problem with the ePPN approach is if the institution re-assigns their internal user identifiers over time (e.g., after a user departs the institution, their identifier is recycled). In this case an ePPN today may not refer to the same user at some time in the future. A more complete discussion of this issue can be found in [4].

The ePTID approach does not suffer from this problem, as an identifier is defined never be reused and hence it will always refer to the same user. The downside of the ePTID approach is that to ensure uniqueness, ePTID must be either computed or retrieved from some persistent storage at the time of use. Both of these approaches created additional infrastructure complexity. Hence many organizations instead choose to adopt policies changes to make ePPNs such that they are not re-assigned (e.g., they do not reassign identifiers even after users depart).

B.1.2 Metadata

InCommon maintains information about its participants and their service deployments that all participants require in order to interact with each other. This information is referred to as “metadata” [56]. All participants will need to initially install InCommon’s metadata and then, typically, run an automated process to maintain a local copy of the most recent metadata to reflect changes in InCommon membership and service information.

B.1.3 Joining InCommon

The steps to joining InCommon are documented on the InCommon website [53]. From a technical perspective, the main steps are:

- Selecting an Administrator and having that person vetted via phone by InCommon. The Administrator should be authoritative for the technical data submitted to InCommon by the organization and is typically a member of the senior technical staff.
- Completing the Participant Operating Agreement [39]. This document needs to be completed by a person or persons familiar with both the technical and policy aspects of the organization’s identity management system and authorized to sign on behalf of the institution.
- Registering the deployment using the InCommon administrative interface [2] so that site information is entered into the InCommon metadata.
- Deploying Shibboleth services, integrating them with the local identity management system or application service(s) in the process. Downloading the InCommon Metadata [38,56] and configuring Shibboleth-enabled services to utilize it [41].

B.1.4 User Support

Like any other service provided by an institution, a user support plan should be in place to help users who encounter difficulties. One aspect of federated identity is that issues can easily span multiple organizations. Because of this, institutions will want to at least be aware of the support points of contact at other key organizations and ideally establish working relationships with them to help debug user issues when they arise.

A challenge particular to NSF CI and federated identity is that it is not unusual for support staff not to have access to the NSF CI due to NSF CI tending to use identity-based access control. Ideally CI projects should allow for access by identity provider support staff to allow that staff to be familiar with the access modality and to aid in debugging.

B.1.5 Computer Security Incident Response

Federated identity presents a new challenge to computer security incident response in that it extends the impact of user credentials being used illicitly by third parties from being a purely localized incident at identity providers to incidents that effect service providers relying on those identity providers. We highly recommend that both identity and service providers incorporate this into their risk assessment processes. We also recommend that organizations ensure that their team responsible for computer security incident response be aware of the possibility of illicitly-used credentials being used through the federated identity system, and incorporate a check for such activity into their incident response process, contacting affected organizations in the event such activity is determined to have taken place.

NSF CI projects are frequently, due to their use of sensitive resources and/or data, more interested in computer security incident response than are typical service providers.

B.2 Technical Deployment for Institutions (Identity Providers)

In this section we provide guidance for the technical aspects of Shibboleth deployment, InCommon membership and supporting NSF CI for institutions representing users which are acting as Identity Providers (IdPs). The majority of these steps are generic to any InCommon identity provider; hence this document summarizes and provides references for the relevant Shibboleth and InCommon documentation, and instead focuses on aspects particularly important to supporting NSF CI.

This section focuses on an institution that is deploying its own Shibboleth services. Alternatives, such as an outsourced deployment, are discussed in Section A.5.

B.2.1 Prerequisite Identity Management System

As discussed in Section A.4.1, federated identity builds on an existing identity management system. While establishing an identity management system is outside the scope of this document, some resources for doing so can be found in Section F.2.

From a technical deployment perspective, a mature identity management system means providing:

- *A well-defined authentication interface*. The Shibboleth IdP software is deployed as protected web application and requires an authentication service, such as Kerberos, LDAP, etc., that can be integrated into a web hosting container to provide authentication.
- *A well-defined attribute interface.* The Shibboleth IdP retrieves user attributes for transport to service providers as discussed in Section A.2.

More details on how these services are used by the IdP are provided in the following section on deploying the IdP software.

B.2.2 Shibboleth Identity Provider Service Deployment

A complete list of Shibboleth deployment steps can be found in the Shibboleth deployment checklist [80] and greater detail on how to perform each of these steps can be found in the Shibboleth support documentation [89], in particular the Shibboleth Getting Started Guide [81] and the Technical Deployers Info Center [86]. Technical details are accurate with version 2.2 of the Shibboleth IdP software, the most recent at the time of this writing.

B.2.2.1 Deploy the Shibboleth Identity Provider Software

Building on the identity management system, the first step is to deploy an appropriate hosting container, typically Apache Tomcat, and the Shibboleth identity provider (IdP) software. Full details can be found in the Shibboleth IdP install guide [21].

As part of this process the deployer will integrate the IdP with the local authentication and attribute services [24]. For authentication, the Shibboleth IdP will be similar to any authenticated web application in that it will need to be configured to interact with the organization's authentication service. Attributes are made available by configuring (or developing for unsupported interfaces) appropriate data connectors [18].

Configuring one or more methods of releasing a persistent identifier as described in Section B.1.1 should be done to support NSF CI.

B.2.2.2 Establishing Auditing

The identity provider administrator should ensure auditing is configured and functional [22] to support debugging, security incident response and gathering usage statistics for planning. Auditing tends to be more important with NSF CI than with other service providers because of what is typically a strong interest in user support and security incident response (as discussed in Section B.1.5). Hence a key goal of auditing would be to identify a user given a report containing information available to a service provider.

B.2.2.3 Joining InCommon and Configuration Metadata Maintenance

The next step would be joining InCommon and configuring metadata as discussed in Section B.1.3. The process of joining InCommon enters the organization's information into the InCommon metadata. The organization then needs to obtain InCommon's metadata [56] so that it can interact with other InCommon participants.

Subsequent to the initial metadata configuration, InCommon will regularly have membership and contact information changes that result in metadata changes. An IdP needs to keep its local copy of the metadata up to date to track these changes. This can be accomplished by configuring the IdP [38] to use a metadata provider that downloads the metadata automatically (e.g., FileBackedHTTPMetadataProvider [23]) or regularly update a local metadata copy with, e.g., cron.

B.2.2.4 Configuring Attribute Release

As discussed in Section B.1.1, a Shibboleth IdP administrator needs to configure attribute release policies so that service providers receive the attributes they require. The organization should determine a process for determining the attribute release policies (see Section C.3.4) and the administrator should implement an initial configuration [19].

At this point an organization would be capable of testing its deployment with other InCommon participants.

B.2.2.5 Replicated Deployment

While load does not tend to be a factor requiring replication, many organizations, when deploying a Shibboleth IdP in production, choose to replicate the identity provider service for reliability. The Shibboleth project provides guidance for such replication [20].

B.2.3 Maintenance

There are a number of ongoing technical maintenance tasks associated with an identity provider deployment. Please see Section C.2.1.7 for a discussion. None tend to be particular to supporting NSF CI.

B.3 Technical Deployment for Cyberinfrastructure Projects (Service Providers)

In this section we turn to technical deployment advice for NSF CI projects acting in the role of service providers, that is, consumers of identities provided by campuses and other institutions acting as identity providers. This whole section regards NSF CI projects and is not highlighted past this paragraph.

In general, CI projects will face a subset of the following challenges in enabling researcher access by InCommon:

1. Integrating the methods their users use to access the project's CI with the web-based profiles supported by InCommon. There are two factors that influence the best solution for how the project interfaces with InCommon:
   - Usage modality, that is, whether users utilize a web browser or command line client to access the project?
   - Authentication method, that is, do users utilize public key infrastructure (PKI) credentials [95], also referred to as "grid certificates", for authentication or some other means?
2. Integration of federated identities with the project identity management system. While federated identity allows projects to rely on identity providers to authenticate their users, the projects are still responsible for determining what privileges (if any) the user possesses within the project, so this portion of the identity management system remains the project's responsibility and must be interconnected with Shibboleth and InCommon by the project.
3. As with any other service provider, undergoing the "boarding process": establishing their attribute needs and arranging attribute release from the identity providers representing their users.
4. Making arrangements for access by members of their user community whose institutions are not currently participating identity providers in InCommon.

This section starts with a brief discussion of PKI Credentials and CILogon, an online service designed to bridge from InCommon to PKI credentials that are commonly used in NSF projects. It then proceeds to discuss each of the challenges listed above and concludes with other issues.

B.3.1 Public Key Infrastructure Credentials and CILogon

It is common for NSF CI projects to use public key infrastructure (PKI) credentials ("grid certificates") for authentication [95]. The use of PKI credentials is common for "grid" command-line clients (e.g., GSI-OpenSSH, GridFTP, GRAM, Condor-G). PKI can be integrated into web portals allowing researchers to authenticate with a username and password, and a PKI credential is obtained for the researcher, for example, from MyProxy [3]. The credential is then used by the portal with a grid client to access PKI-enabled services on the researcher's behalf.

The CILogon Service [8] is a NSF-funded service to bridge between InCommon and CI that utilizes PKI credentials. CILogon can either deliver a PKI credential to the user's local system or to a project web portal. In typical usage, a CI project portal would redirect a user to the CILogon service, which would authenticate the user utilizing InCommon, generate an X.509 credential as a result of that authentication and then securely pass that credential to the project portal (details of how this is done are available at [7]). This credential serves both to establish the user's identity for the portal and can be used by the portal to access other services on the user's behalf (described subsequently in Section B.3.2.2.4).

B.3.2 CI Project InCommon Solutions

Table 2 shows the solutions available based on the following two factors discussed in the introduction to this section:

- The project's usage modality: does the project support access via a web-based interface or a command-line application (or other non-web interface such as a programmatic API)?
- The project's authentication mechanism: does the project support access via PKI, or other mechanisms?

Table 2: Solutions depending on project's normal mode of access and authentication mechanism.

Usage Modality	Authentication Mechanism	
	PKI	Other
Web-based	CILogon with project portal	Shibboleth-protected portal
Command-line	CILogon with PKI-enabled command line clients	No current solution available

The solutions are not mutually exclusive; projects may want to deploy more than one if they support multiple usage methods – for example, web and SSH access. The four solutions are summarized in the following list and described in detail in the following subsections:

1. Projects providing a web interface and not using PKI can deploy the standard Shibboleth Service Provider (SP) software to Shibboleth-enable their web interface and then join InCommon as would be normal for an InCommon service provider.
2. Projects providing a web interface and using PKI credentials (e.g., projects using MyProxy) can utilize the CILogon service to authenticate the users via InCommon and deliver a PKI credential to the project portal for the user.

3. Projects providing a command line interface and using PKI credentials can utilize the CILogon service, but, unlike the previous scenario, have the CILogon service deliver a PKI credential to the user's local system for use by PKI-enabled applications (e.g., GSI-SSH, GridFTP).
4. Projects that are current utilizing a command-line interface and authentication other than PKI currently have no good solution available to them. The only guidance this document can give is that the project transition to one of the other scenarios or monitor the items discussed in the future work section (F.1), namely MoonShot and the Federated SSH work.

Some examples of projects utilizing or exploring these options at the time of this writing, which may have experiences to share, are:

- TeraGrid [4] utilizes a variant of solution (3). It's solution was implemented as a processor of CILogon. TeraGrid is in process of integrating Shibboleth support into the TeraGrid User Portal [92] to support solution (2) in addition.
- InCommon access to research.gov is being piloted by NSF [58], representing an implementation of solution (1).
- The Indiana Clinical and Translational Sciences Institute [43] provides for InCommon-based access to its web site as an implementation of solution (1).
- The Open Science Grid [70], DataONE [12] and Ocean Observatory Initiative [69] are in process of exploring or implementing a CILogon-based approach – (2) and/or (3).

B.3.2.1 Shibboleth-protecting a Web Portal

For projects that utilize a web portal as their user interface, deploying the Shibboleth SP software to Shibboleth-enable that web portal is an option. This is done as is typical with any Shibboleth SP deployment; hence we summarize the steps here calling out issues particular to NSF CI.

As with an identity provider deployment, it is recommended that this be undertaken with a prototype deployment first and then transitioned to a production portal.

Note that a major challenge to this approach is arranging attribute release from all the identity providers who represent the project's users as discussed in Section B.1.1.

B.3.2.1.1 Deploying the Shibboleth SP Software

The first step is to deploy the Shibboleth SP software [82] to Shibboleth-enable the project web portal. How challenging this will be depends on what technology is in use to host the portal and how suited the application is itself to having authentication performed outside the application.

In terms of hosting platforms, the Shibboleth SP software works well with the Apache HTTPd and Microsoft IIS platforms, and documentation also exists to couple it with Java-based containers (e.g., Tomcat) [45]. Outside of these technologies you are more likely to find challenges. The best advice is to try and find via, for example, the Shibboleth users email list or a web search engine, someone else who has undertaken Shibboleth integration with your particular technology. Undertaking integration with a technology for the first time is likely to be a significant challenge.

The level of effort to modify the application to be Shibboleth-protected will vary depending on whether the software was written with modular authentication in mind. Many services have a 'baked in' identity management solution and modifying the software to support federated identity can be significant effort. Much research software, developed as research itself by computer scientists or informaticians, may have no concept of security built in at all, which is actually easier to integrate, as coarse-grained access control lists can be implemented by the container and the application unmodified. The Internet2 wiki maintains a page with services and applications known to work well with Shibboleth [76].

B.3.2.1.2 Joining InCommon

A NSF CI project may join InCommon itself or become a service provider under the auspices of an existing InCommon member. Please see Section C.4.2 for a discussion.

If the NSF CI project joins InCommon itself, the process is very similar as the process described for identity providers in Section B.1.3, namely selecting an Administrator and having them vetted, completing the Participant Operating Agreement, registering the site's configuration with InCommon, and installing the InCommon metadata.

B.3.2.1.3 Arranging Attribute Release

Since InCommon does not dictate that identity providers release any set of attributes to other InCommon members or provide any metadata exposing attribute release policies of members, after registering their service provider in InCommon, the project needs to contact the identity providers of its users and arrange for attribute release as described in Section B.1.1.

This is unfortunately a time-consuming manual process, and subsequently making additions to this list of attributes will require re-contacting the identity providers. Hence it is strongly suggested that the project ensure they understand their requirements in this regard before undertaking this task.

A discussion of the attributes commonly required is found in Section C.4.3. Typically these attributes are used to map to a user's entry in a local identity database as described subsequently in Section B.3.3.

Note that attribute release policies are written to release attributes to a specific service provider identifier, which means that changes to a service provider identifier are very painful, as they require contacting all identity providers to arrange the change of service provider identifier.

B.3.2.1.4 Maintenance

There are several components of the service provider deployment that require ongoing maintenance, which are very similar to the maintenance for an identity provider as described in section B.2.3:

- *InCommon Metadata*: InCommon will regularly have membership changes and contact information for existing members may also change from time to time. These changes will be reflected as changes in the metadata. Deployers can either configure a service provider [38,41] to use one of the metadata providers that download the metadata automatically (e.g., FileBackedHTTPMetadata-Provider) or regularly update the local metadata with, for example, cron.
- *Local Metadata Information*: Changes in the local deployment configuration may result in changes to the institutions metadata, which will need to be communicated to InCommon so that InCommon metadata remains up-to-date with regards to the local organization.
- *Software*: As with any software, the Shibboleth SP software needs to be maintained with bug and security fixes. The appropriate announcement lists [83] should be monitored so that the organization is cognizant of fixes and new versions, and can arrange upgrades as appropriate.
- *InCommon POP*: If processes or policies for the local organization change, these will need to be reflected in the organization's participant operational practices.
- *Adding Supported Identity Providers*: Each IdP may configure the attributes they release slightly differently and this needs to be configured in the SP attribute configuration so that those attributes are exposed to the application logic appropriately.
- *Audit Log Rotation*: As is typical with logs, rotation and retention policies should be defined and implemented. We recommend retaining logs for at least three months to accommodate security incidents that may not be immediately detected, though we acknowledge the exact amount of retention will depend on the size of available storage, the amount of SP usage, and any relevant policies.
- *Reliability and Scalability*: If replication of the service provider is needed for load-balancing or reliability, details for doing so may be found on the Internet2 web site [60].

B.3.2.1.5 Accessing Other Services (the n-Tier Problem)

A common workflow is for a user to interact with a project portal and for that portal to then act on behalf of the user to coordinate other resources. For example, the portal may access data stored on another service for processing. This is often refereed to as the "n-tier" problem, with the portal representing the first tier of user interaction and then services coordinated by the portal as a second tier, and further services would represent the third tier and so forth.

Extensions to Shibboleth to support this use case [10] are still in the pipeline and not available for deployment. Currently the only option with this approach is to establish service-level trust between the project portal and 2nd tier services. The 2nd tier services will explicitly need to trust the project portal to be acting properly on behalf of a user and not require any proof that this is the case. The portal will typically authenticate using a credential issued to the portal specifically for this purpose. An example of this is the TeraGrid Science Gateway security model [75].

B.3.2.2 Using CILogon with a Project Web Portal

We now turn to the second solution, using CILogon to allow InCommon access through a project web portal. In this deployment scenario the project, rather than joining InCommon itself, establishes a relationship with the CILogon Service. Users are redirected to CILogon, which provides a PKI credential to the project portal as a result of InCommon authentication. This credential serves both to establish the user's identifier for the portal and can be used by the portal to access other services on the user's behalf (described subsequently in Section B.3.2.2.4). The text in this section provides a high-level overview of CILogon interaction, for details please see [7].

B.3.2.2.1 Integrating Support for the CILogon Service into the Project Portal

The CILogon service uses the OAuth protocol [68] to coordinate the authentication of the user and delegation of the resulting user credential to the project portal. The portal will need to integrate CILogon client code in order to handle its side of this process.

B.3.2.2.2 Establishing A Relationship with the CILogon Service

In addition to integration of client code, the CILogon service requires an exchange of cryptographic material with the project portal to allow for subsequent secured communication. A project portal administrator needs to contact the CILogon project to arrange this exchange of cryptographic information.

B.3.2.2.3 Effort and Maintenance

The level of effort to integrate with CILogon will vary depending on the portal implementation, but is roughly equivalent to integrating a new authentication mechanism.

Maintenance is required to keep client software and configuration for the relationship with CILogon up to date. As with any software and service, updates will be needed periodically to address bugs and security issues. The portal administrator should ensure they are on appropriate CILogon email lists in order to be cognizant of needed changes.

B.3.2.2.4 Accessing Other Services (the n-Tier Problem)

As described in Section B.3.2.1.5, some project portals may want to support workflows in which the portal acts on the user's behalf to coordinate other services. In addition to the explicit trust model described in Section B.3.2.1.5, the PKI credential received from the CILogon Service allows the project portal to coordinate on the user's behalf any service that accepts the user's PKI credential.

B.3.2.3 Using CILogon with Command Line Clients

We now turn to the third solution, utilizing command-line clients with PKI credentials. In this deployment scenario, the project provides services that are accessed by command line applications or other clients that run on the user's computer and authenticate via a PKI credential stored on their computer. In this scenario the delivery of the PKI credential to the user's computer is done entirely by the CILogon Service, the project would direct the user to the CILogon Service to initiate this process. The project would need to provide a mechanism for the user to register the identifier in the PKI credential with the project so that they recognize the credential when it is presented. Since this is a one-time, or at least infrequent, event, a typical mechanism is to have the user authenticate via an existing mechanism and provide the PKI identifier (e.g., this is how TeraGrid does it [4]), or to have the user contact a member of the project staff and be manually vetted (e.g., this is how the Open Science Grid does it [71]).

B.3.2.3.1 Effort and Maintenance

Establishing the ability to trust PKI credentials issued by the CILogon service is equivalent to establishing trust with any other PKI certificate issuer (i.e., certificate authority), which is typically not a difficult technical challenge. Developing the mechanism to allow users to bind their CILogon identity to their existing project identity requires developing a secured application capable of modifying a protected database, roughly a two-week task by a developer familiar with the services based on the authors' experiences, plus appropriate time for testing and deployment.

B.3.3 Integration with Project Identity Management System

We turn now to issues common across all three solutions covered in the previous subsections. Federated identity allows the service provider to outsource authentication of the user to the identity provider, though typically the service provider will still need to maintain access control information about the user (e.g., what privileges the user has within the scope of the CI project).

Based on the experiences of the authors [4], it is suggested that the project maintain a project-internal identifier for the user, to which they map the access control information as well as identifier(s) provided by the user's identity provider(s). This is recommended for several reasons:

- *Supporting multiple authentication methods*. CI projects with previous or alternate authentication methods (e.g., OpenID, username and password, SSH keys) can map the identifiers of those methods to this internal identifier.

- *Supporting user movement between institutions*. Since the federated identity fabric does not support the portability of a user identity across organizations, if a user moves from one organization to another, their identifier will change. By allowing for a mapping, a CI project can remap the new identifier to the internal identifier and not require any other re-enrollment of the user.
- *Supporting joint appointments*. Allowing multiple external identifiers to be mapped to the internal identifier addresses users who may present identities from different institutions at different times.
- *Auditing*. The project can track when the binding was made and, for institutions that potentially re-issue identifiers as discussed in Section B.2.1, re-affirm the binding on a regular basis.
- *Staged Adoption*. The project can migrate users from existing authentication systems to InCommon at whatever pace they desire.

The result of this approach is that it allows the process to establish a binding between the InCommon identifier (or its equivalent from the CILogon service) and the existing project identifier and subsequently map the user's InCommon identifier to the project identifier.

B.3.3.1 Supporting Multiple Authentication Mechanisms

To continue the Jean Blue story from Section A.3.1:

> "While at her collaboration platform, Jean notes that some content is now six months old and is now eligible for public access, though her funding agency requires to know the number of users of the data. She sets the access policies on this data to be world readable, but requiring either a federated identity or a social identity be used to access the data, allowing her to provide usage statistics."

While federated identity and InCommon are important tools within the scope of authentication and authorization, they do not serve all the needs of either the institutions or the virtual organizations. Many projects will have existing authentication mechanisms and users may already have accounts elsewhere – at social networking sites such as Google, MSN, Facebook, etc. – that a project may want to leverage. (We include a discussion of social networking as an alternative in Section A.5).

One advantage of the approach where the project maintains an internal identifier for its users and maps external identities to that internal identifier is that multiple external identifiers from multiple mechanisms can be mapped to that internal identifier. This allows a project to support however many authentication mechanisms it sees fit. For example, this was used in the TeraGrid to add federated identity support to existing PKI, SSH RSA keys, one-time passwords, and standard passwords [4].

B.3.4 User Support and Security Incident Response

Federated identity brings some additional user support challenges since user access now involves services outside of those controlled by the project. It is recommended that projects do the following to aid in solving user support problems:

- Establish points of contact for support staff at services representing users, namely identity providers in InCommon or staff at the CILogon Service.
- Establish points of contact for security incident response at services representing users, namely identity providers in InCommon or staff at the CILogon Service.
- Establish a simple service that is robust as possible and tests basic user authentication and attribute release to determine if problems are related to federated identity. For example, for projects with a Shibboleth-enabled web portal this would validate receipt of the project's required attributes.
- For the service in the previous bullet, establish a procedure allowing support staff at services representing users the means to use it for testing. This helps establish a means for end-to-end testing that doesn't require user involvement.

B.3.5 Auditing and Logging

Auditing of federated identity is recommended for both user support and security incident response. With regards to user support, user access attempts could fail because they are utilizing unrecognized identity providers, identity providers that are not releasing required attributes, errors in configuration, network failures or software flaws. Ideally, access mechanisms will detect these failures and provide feedback to users as they happen, but audit logs should allow for debugging.

For security incident response, in the event a user interaction is suspected to be the work of an imposter, it is important that the project be able to identify what identity provider and user identifier were involved as well as other pertinent information (e.g., IP addresses of computers from which the client activity originated).

Additionally, projects may find aggregated usage information useful for reporting purposes.

B.3.6 Supporting Users Not at InCommon Institutions

To continue the story of Jean Blue:

> "Jean then pursues one of her action items, to add several colleagues from around the world to the collaboration. She goes to her remote collaboration management platform, transparently authenticated using her earlier institutional login. Once there, she adds the email addresses of each of these new colleagues, to her group. She then clicks invite. Each of the collaborators instantly receives an invite email. By clicking on the URL within, they are prompted to authenticate to their own home institutions and then have their identity added to the collaboration

> group, instantly allowing access to the group wiki, adding them to the group email list, permitting them to schedule videoconferences, etc. Through federation, local authentication, wherever in the world, can be used to access global research resources."

A problem that a project may run into is a user who is not located at an InCommon participating institution, either because that institution simply has not joined InCommon yet, or perhaps they are ineligible because they are outside the U.S.

In this circumstance, there are several solutions that can be explored:

1. Research and Education federations for identity have been developing in many countries, particularly within those countries that have traditionally had extensive research relationships with the US, for example [74]. In some of these countries, coverage of the R&E community is now 100% and there is a robust federation infrastructure underlying almost all inter-institutional relationships. If the user's institution acts as an identity provider as part of another compatible federation, for example, one of the European federations, the CI project could, assuming the Federation's policies allow it, join that federation. Such membership in multiple federations, while not something the community has a lot of experience with, is possible at least in theory.

   Alternatively, inter-federation between the federations could be explored. Activities are underway to "inter-federate", that is to connect federations together at peering points to allow credentials to flow throughout the international research and education community. At this time, this would be a more experimental, time-consuming course.

2. If the user's institution runs a compatible identity provider, the CI project can bilaterally peer with that institution [79,85]. The effort for doing this technically is roughly equivalent to joining a federation, though the parties would need to create their own mechanisms for maintaining up-to-date metadata between themselves. From a policy perspective, this process can be as formal or informal as the two parties desire.

3. The project can explore obtaining a guest account for the user at an InCommon participant. Typically this entails the project having a relationship with the participating institutions (e.g., the project PI is a faculty member).

4. Users can use a free identity provider that is an InCommon participant (e.g., ProtectNetwork [73]) to access InCommon-protected services. For projects using CILogon, it supports ProtectNetwork for just this reason. This approach is perhaps easiest for the CI project, but places the most burden on the users. It also means that, as with use of non-SAML identities discussed in the next bullet, any attributes will be self-asserted by the user and have a lower level of trust than those asserted by a institution.

5. The project can support the use of identities used by Facebook, Google, Yahoo, etc. (typically through the OpenID or OAUTH protocols) and have the user authenticate through one of those services. For example, the CILogon

service [8] supports both InCommon and OpenID. A full discussion of accomplishing this is beyond the scope of this document, however, following the advice in Section B.3.3 with regards to supporting multiple authentication methods should prove to be a large aid in accomplishing such support. Projects should keep in mind that these identities tend to be self-asserted and have lower levels of trust.

B.3.7 Provisioning

A decision to be made by the CI project is how new users will be enrolled in the project's identity management system. The challenge is that the user's identity, asserted by their identity provider, is not predictable by the project.

The emerging best practice for enrollment in the context of federated identity is to design an InCommon-protected interface for the user to request enrollment. The interface requests from the user whatever attributes the CI project requires to complete enrollment (either via the federated identity mechanism or via manual web form completion). The CI project can then vet the enrollment request as they see fit. (This interface can be the same interface discussed in Section B.3.3 to map the user's federated identity to an existing identity by requesting a second authentication of the existing identity.)

A Roadmap for using NSF Cyberinfrastructure with InCommon

Guide to Policy and Business Processes

Abstract

The *Guide to Policy and Business Processes* is intended for managers and policy makers, and is a guide regarding the policy, privacy, financial and other concerns for InCommon deployment. Again it is both for staff from campuses (and other organizations which serve as a home to scientists and engineers) and NSF cyberinfrastructure projects.

C Guide to Policy and Business Processes for Deployment

This section is meant to provide guidance with regards to policy and business process issues involved in the adoption of Shibboleth and joining InCommon. It is intended for policy decision makers in institutions and CI projects. This section starts with a brief introduction to the policy and business process issues, and follows with two subsections, one for institutions (identity providers) and one for CI projects (service providers).

For the most part, issues are not specific to supporting NSF CI, but have some details that are influenced by that support. Those details related to NSF CI are highlighted in the same manner as this paragraph..

C.1 Introduction to Policy and Business Process Issues

In this section we provide a brief introduction to policy and business process issues common to both institutions representing users and CI project providing services to those users.

C.1.1 InCommon Participant Agreements and Operating Practices

InCommon members are expected to complete two documents. The first is the InCommon participant agreement [39], a legal document laying out the contractual relationship between the participant and InCommon. This agreement needs to undergo whatever process the institution has for signing contracts, which can be time consuming and thus should be initiated early in the process.

NSF CI projects typically are not legal entities in their own right and will need to have an organization represent the project in the process to sign the agreement. Typically this would be the lead institution for the project.

The second document is the Participant Operating Practices [39], which describes the technical, policy and business practices the organization or project has in place for identity management, attribute use, privacy, etc. This document covers both existing identity management practices and practices related to federated identity. This document is completed in good faith and does not require a signature.

C.1.2 User Identifiers

For those used to traditional authentication systems centered around usernames, one slightly counterintuitive aspect of Shibboleth is that what one would typically think of as a user identifier is conveyed as an attribute. The reason is due to Shibboleth allowing for different types of identifiers to be conveyed in different situations to support privacy and anonymous access. Such identifiers are often needed by NSF CI projects and will need to be considered by those projects when they determine their attribute requirements.

In order to support persistent identifiers needed by NSF CI, institutions need to understand their identity reuse policies. For example, if a faculty departs the institution, will they institution re-assign that identity to another incoming user? If so, after how long?

C.1.3 Attribute Release and Consumption

A strength of Shibboleth is its ability to release attributes in a controlled manner. For each service provider, the Shibboleth identity provider administrator can configure what attributes will be released. Future technology, such as uApprove, discussed in Section F.1, would put the users in the loop and give them the opportunity to consent or deny the release.

Service providers consume attributes released to them to perform access control, personalization, etc. InCommon makes no requirement of its participants for what attributes they release to other InCommon participants. While an institution could choose to release attributes freely, most choose not to do so due to policy (e.g., FERPA) and privacy concerns, and instead establish a policy for deciding what attributes will be released to what services. In practice, this means that service provides must request the release of attributes they require from institutions representing their users.

C.1.4 Levels of Assurance

A “level of assurance” with respect to identity management is used convey the amount of trust one can have in an asserted identity. This is a complicated issue, covering, among other things, vetting practices of the institution making the assertion, the technical specifics of the authentication process, and institutional policies regarding password changing. For more detail, the reader is directed to NIST Special Publication 800-63 [62].

Strictly speaking, normal participation in InCommon by an identity provider guarantees nothing in terms of level of assurance; while an identity provider is expected to document their practices and policies in the Participant Operating Practices, there is no requirements on what those practices and polices are, or any requirements that they be audited. InCommon supports two levels of assurances that members may voluntarily achieve: Bronze and Silver [35], which are equivalent to levels 1 and 2 respectively in NIST Special Publication 800-63. Very simply put, Bronze entails basic authentication and identity management practices one could reasonably expect for a university with a mature identity management system to have in place, while Silver entails stronger vetting, authentication, and institutional audit that are more typical for sensitive applications (e.g., access to sensitive medical data).

At the time of this writing, InCommon Silver is still in the very early stages of adoption with only a small number of campuses in process of achieving it. Readers interested in exploring what achieving Silver entails, may be interested in the CIC report exploring Silver adoption [6].

This document goes into no further detail on the matter except to say that CI projects providing access to sensitive data or resources may want to examine the

Bronze and Silver profiles, and decide if requiring one or the other would benefit the project. Projects should be aware that, currently, the limited adoption of these profiles by InCommon participants will limit InCommon's usefulness to the project. Experience in the TeraGrid project [4] has shown that identity vetting done by the project may augment lower levels of assurance to provide a higher effective level of assurance to the project.

C.2 Effort Required for Shibboleth Deployment and InCommon Membership

In Section A.4.4 we provided a table summarizing effort to join and maintain InCommon membership. We expand on that table here with more detail on each task.

C.2.1.1 Leadership

The entirety of the process of deploying federated identity and joining InCommon will require leadership from someone who is capable of orchestrating activities across technical, policy and business process units. On a campus, this typically requires someone at the level of the CIO's office, or a delegate, to lead the deployment. Support of other campus leadership and administrative units (the Registrars or Human Resources offices are often important) is crucial. Coordination across the units to provide support and outreach to the campus researchers, students and other users will be important to achieve successful adoption and support.

For a CI project, it will require analogous leadership by the project's technical lead with support from other project leadership.

C.2.1.2 Policy and Business Process Issues

From a policy and business process perspective, joining InCommon requires documenting existing identity management practices, and establishing policies for identity release or identity information use, for a CI project. This effort is equivalent to an effort to change a substantial authentication mechanism, for example, changing a policy on password strength. It requires mastery of current practices to understand who is impacted, bringing those parties together, reaching agreement on the change, and then implementing the new policies.

C.2.1.3 Signing the InCommon Participant Contract

Joining InCommon requires signing the Participant Agreement [39]. Signing the Participant Agreement will require engaging the organization's or project's usual process for contract signing.

C.2.1.4 Technical Effort to Deploy a Shibboleth Identity Provider

Deploying an identity provider system is roughly equivalent in terms of effort to deploying a web single sign-on technology (e.g., CAS [5]). It requires deploying a secured web service that acts as the authentication service, and establishing new support, monitor, backup and fail-over processes for that service. As with any such service, it is typical for organizations to deploy it as a prototype first and then move to a production-level service over time. Hardware requirements are modest and load on the service is not a typical problem reported by organizations. It is, however, common for an organization to replicate a production identity provider service for reliability.

C.2.1.5 Technical Effort to Deploy a Shibboleth Service Provider

Federated identity by itself is just a foundation on which collaboration infrastructures can be built, and by itself is not useful to researchers; federated identity needs to be integrated into services such as web sites or applications to become useful. CI projects will need to deploy one or more service providers and campuses will probably also choose to do so in order to use federated identity internally.

Deploying a Shibboleth-protected service is roughly equivalent in effort to deploying any single sign-on protected web application, that is, the effort varies widely depending whether the hosting container is known to work with Shibboleth and whether the application has modular identity management that lends it to easy integration with federated identity. A more detailed discussion can be found in Section B.3.2.1.1 of the Technical Guide.

C.2.1.6 Effort Required to Add Federated Partners

Once an InCommon federated identity infrastructure is deployed, a common activity is adding a federated partner. Such partners can be identity providers representing users with whom the organization wishes to collaborate or provide services, or service providers representing CI or other services to which an organization wishes to provide their users access. Such additions may be initiated by the identity provider (e.g., by researchers or business needs) or by the service provider (e.g., a CI project wanting to establish a partnership for current or anticipated researchers).

From the technical point of view, InCommon and Shibboleth make the addition of such partners straightforward by defining the protocol format and privacy controls. This means such additions amount to configuration changes to configure attribute release and use for identity and service providers, respectively.

InCommon, however, does not define policies in regards to the release of attributes among its members, meaning those policies must be negotiated between interacting pairs of partners directly[2]. In practice this means the identity provider, taking into account their local privacy and policy considerations, deciding what user attributes they are comfortable exposing to a service provider, given that service provider's needs and policies. An identity provider, by defining up front who is authorized to make this decision and what its requirements are in terms of what it allows and expects, can greatly streamline this process. Similarly, a service provider, by defining up front what identity information it requires and how it will use that information, can make the decision straightforward for identity providers.

[2] The goal of uApprove [93] is to change attribute release from an institutional policy to a decision by the user at time of their access to services. This process is expected to undergo significant changes when an identity provider adopts uApprove.

C.2.1.7 Service Maintenance

The ongoing effort for maintaining InCommon and Shibboleth is comparable to maintaining any web single sign-on system, with requirements on maintaining software, configuration, certificates and fielding user support requests:

- *Software maintenance.* As with any software service, the software will periodically need to be upgraded due to security and compatibility issues. The appropriate announcement email lists should be monitored so that the organization is cognizant of new versions and can arrange upgrades as appropriate.
- *Federation configuration maintenance.* InCommon will regularly have membership changes and contact information for existing members may also change from time to time. This information is reflected in InCommon's configuration data, commonly called "metadata". The distribution of metadata is normally an automated process requiring no manual attention. Changes in the local configuration may result in changes to an institution's metadata, which will need to be communicated to InCommon so that InCommon metadata remains up-to-date.
- *InCommon Participant Operating Practices*: If processes or policies for the local organization change, these will need to be reflected in the organization's participant operational practices.
- *Attribute Release Configuration.* As discussed throughout this document, Shibboleth identity providers have configuration limiting what user attributes are released to what service providers. Unless a highly permissive policy is adopted, that configuration will need to be modified as service providers are added. This is discussed in Section C.3.4. For service providers, if their attribute requirements change, they will need to contact relevant identity providers to request configuration changes.
- *Certificates.* Identity and service providers require a SSL server certificate for their user-facing web applications as with any other secured web server. Additionally, these services require a certificate for their Shibboleth endpoints, however these certificates can be (and are recommended to be) self-signed and renewed every three years [97].
- *User support requests.* As with any web single sign-on service there will be user support requests. Federation adds complexity in that issues may be caused by interactions among federated partners. In practice this means maintaining good contact information for support staff at partners to quickly resolve such issues. InCommon metadata has contact information for Shibboleth administration staff, which helps with locating other staff at an organization.
- *Work force.* As with any other service of this complexity, staff with up-to-date training is necessary to maintain it. Level of effort is analogous to staying abreast of other software systems and InCommon offers participants

numerous online and in-person training opportunities as described in Section F.3.2. Internet2 Member Meetings [47] are also a good source of information.

C.3 Institutional Deployment: Policy and Business Process Issues

In this section we discuss policy and business process issues for deploying Shibboleth and joining InCommon for an institution representing NSF scientists and engineers as an identity provider. The equivalent discussion for a CI project follows in Section C.4. These issues are for the most part generic to any identity provider joining InCommon and not specific to supporting NSF CI.

As with the Technical Deployment, and as discussed in Section A.4.1, a prerequisite to federated identity management is having an institutional identity management system in place. In the context of this section, this primarily means understanding the institution's identity management policies and practices in order to complete the InCommon Participant Operating Procedures as discussed previously in Section C.1.1.

C.3.1 Policy Steps to Joining InCommon

Figure 2 gives a high-level, conceptual roadmap of the policy and business steps in joining InCommon. The figure shows a conceptual flow of activities and should not be taken as strict dependency graph; in fact, some of the later steps such as signing the Participant Agreement, should be initiated early in the process. We discuss each of the steps, and its associated issues and implications for supporting NSF CI in the following sections.

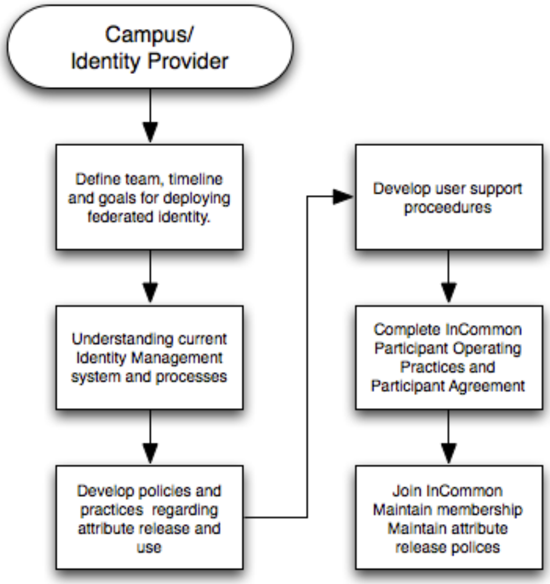
Campus/ Identity Provider
Define team, timeline and goals for deploying federated identity.
Understanding current Identity Management system and processes
Develop policies and practices regarding attribute release and use
Develop user support procedures
Complete InCommon Participant Operating Practices and Participant Agreement
Join InCommon Maintain membership Maintain attribute release polices

Figure 2: High-level roadmap of policies issues for a campus (identity provider) joining InCommon.

C.3.2 Project Planning

As with any complex project, the institution should undertake appropriate project planning. This document is not meant to replace a project plan. Some suggested reading and steps for this planning process includes:

- Determine InCommon eligibility. This is typically not an issue for universities. Details can be found in Section C.4.2.
- Identify institutional need and assess the value to the research and educational enterprise. Identify some key projects that will demonstrate that value.
- Estimate level of effort as described in Sections A.4.2 and C.2.
- Explore alternatives to Shibboleth and InCommon (see Section A.5).

- Review the numerous resources InCommon has available for planning a federated identity deployment; a list as of the time of writing this document can be found in Section F.3.

It is typical to plan on setting up a prototype Shibboleth deployment for testing and evaluation before executing a production deployment. If an organization takes this path, identifying one ore more NSF CI projects of interest to the institution's researchers and the researchers using those projects to test the prototype is recommended.

C.3.3 Understanding Current Identity Management Polices and Processes

The institution needs to understand its current identity management policies and procedures. The primary reasons for this are to plan the technical integration between Shibboleth and the identity management system as discussed in Section B.2.2 of the Technical Guide, and to complete the InCommon Participant Operations Practices [39], discussed in Section C.1.1. Those with the better-defined identity management policies and practices, including risk management strategies for unintended identity information leakage (see Section C.3.7), will more easily provision a federated identity system through InCommon.

As discussed in Section C.1.2, NSF CI often has a requirement for identifiers for researchers, which are, ideally, not re-assigned to other users, e.g., when a researcher departs the organization. During this study of the existing identity management system, forming a solid understanding of when and under what circumstances a user identifier can be reassigned will be helpful to supporting NSF CI.

C.3.4 Attributed Release Policies

Shibboleth provides configuration to control what user attributes are released to what service providers, including NSF CI projects. Institutions should anticipate requests for attribute release by service provides and have a policy in place for how they should response to help expedite these interactions.

One can roughly categorize these requests for attribute release into two cases: requests for identifying attributes or requests for aggregate attributes. The latter consist of attributes such as the user's role at the institution (e.g., Faculty) while the former consistent of persistent identifiers or other personally identifying information such as name or email address. Typically NSF CI projects will make requests for identifying attributes. These requests tend to be the more challenging from a policy perspective, so we focus on those requests here.

In the experience of the Roadmap authors, the following procedures are commonly used to respond to these attribute release requests:

1. *Highly permissive approach*: Attributes are released to other InCommon members freely or with light scrutiny by operations staff.
2. *Justification required approach*: In this case, a justification for attribute release needs to be presented by the requestor, potentially including a list of

users at the institution in question and intended use of the released identification and attributes. An identified individual or individuals inside the organization evaluate the justification.

3. *Champion required approach*: A person at the institution needs to step forward to champion the service provider's attribute release request. This is typically coupled with a justification as discussed in the previous bullet (2).

This Roadmap strongly advocates against the champion-based approach because users of NSF CI projects tend to be domain scientists, and putting them front and center in discussions regarding federated identity tends to work poorly as they usually are not identity management experts. It is advisable to consult researchers identified in a project's justification to vouch for their use of that project's CI and elaborate on their science needs to ensure those are being served. Our advice therefore is procedure (2), including consulting some set of the researchers involved.

C.3.5 User support Procedures

As with any similar service, user support personnel will need to be trained in the new service. Additionally, because of federated identity's nature of being distributed among multiple organizations, it is suggested than user support personnel establish relationships with their counterpart at service providers of significant interest, or at least know those counterparts' contact information readily.

One challenge that NSF CI often brings as compared to standard InCommon service providers is that institutional support personnel often do not have access themselves to the NSF CI, since that access tends to be granted on a user-by-user basis rather than to subsets of the university population. This strengthens the previous suggestion to establish a relationship with the NSF CI project support staff and also explore if access can be arranged for troubleshooting (e.g., the TeraGrid offers a 'Campus Champions' program [91] which would provide this).

C.3.6 Joining InCommon: Agreement and Documentation

The process of joining InCommon includes signing the Participant Agreement and completing the Participant Operating Practices as described in Section C.1.1. It also includes establishing an InCommon Executive. The Executive will be identified to InCommon as the ultimate authority with regards to the institution's membership. In practice, the Executive assigns an administrative point of contact that will handle day-to-day technical activities. It is recommended that the Executive be a CIO, VP of IT, or similar position in the institution.

C.3.7 Risk Management for Accidental Attribute Release

Despite technical and policy controls, it is possible due to misconfiguration or software flaw for user attributes to be disseminated beyond what is intended. Organizations should consider in their risk profile and/or author policy for such a situation. Questions to consider include: How will the organization handle a misconfiguration event that leads to attributes being released by accident? How will accidental release by a partnered service provider be handled?

C.4 Cyberinfrastructure Deployment: Policy and Other Issues

This section focuses on CI Projects who are joining InCommon.

C.4.1 Policy Steps to Joining InCommon

The major policy steps in joining InCommon for CI project are:

1. *Establish Membership Eligibility* [1] as discussed subsequently in Section C.4.2.
2. *Signing the Participant Agreement* [39] as discussed in C.1.1.
3. *Establish the InCommon Executive*. The Executive will be identified to InCommon as the ultimate authority with regards to the institution's membership. In practice, the Executive assigns an administrative point of contact that will handle day-to-day technical activities. It is recommended that the Executive be a technical lead for the project.
4. *Completing and submitting the InCommon Participant Operating Practices* [39]. This process will often include collecting practices from different units of the institution.
5. *Determine Required Attributes.* Discussed subsequently in C.4.3.
6. *Identify Institutions of Interest and Arrange Attribute Release.* The project needs to identify what institutions represent their users and then contact those institutions to arrange attribute release. How the institutions are identified is project specific; for an existing project it can be accomplished by surveying existing users, while new projects may need to make educated guesses. Arranging attribute release is discussed in the Technical Guide in Section B.3.2.1.3.

C.4.2 InCommon Membership Eligibility

InCommon has requirements for membership [1] that entail members being a U.S. Institution of Higher Education or sponsored by InCommon member. The sponsorship process [40] is not financial and entails having an existing InCommon member request membership on behalf of the project.

The simplest route for an NSF CI project is not to join InCommon as a separate entity, but instead have a relationship with an InCommon member institution (e.g., the project PI's or co-PI's home institution) such that it can be registered as a service provider under that institution's membership. Since InCommon membership allows for the registration of 50 service providers per identity provider [31], most institutions are likely to have such registrations to spare. In this case the project will need to work with the institution's InCommon administrator to update the institution's Participant Operating Practices and metadata.

C.4.3 Determining Federated Identity Attribute Needs

Shibboleth was originally established so service providers could ascertain institutional affiliation of users without sharing any personally identifying information. This privacy-preserving mode of operation works well for services such as online journal subscriptions.

However, service providers providing access to sensitive data or other resources typically require identification of individual users (i.e., authentication). Many NSF CI projects fall into this category. When using Shibboleth, user identifiers of this sort are conveyed to a service provider as attributes, hence although one would normally call this authentication, in the context of Shibboleth, it falls into the attribute release process.

Other attributes tend to fall into either contact information for the user (e.g., email, phone number) or other information about the user that is useful to the project for aggregate reporting to funding agencies (e.g., field of study, professional title, ethnicity, gender). Projects have the choice of collecting this information themselves when enrolling the user (e.g., via a web form) or requesting these attributes via federated identity.

The advantage of collecting the information directly from the user is simplicity. The user's institution need not be involved in the policy decision and privacy concerns are simplified since it is clear the user is meaning to provide the attributes when they fill out the forms. The disadvantages of this approach is the user is trusted to accurately assert the information and collecting the information requires manual action by the user, meaning it causes inconvenience to do so frequently.

The advantages of receiving the attributes regularly via federated identity are that they are more likely to be kept up to date by the user's home institution (this is mainly a factor for contact information such as telephone numbers), they are typically vetted by the institution, providing greater assurance of accuracy, and they are provided automatically by Shibboleth for each user session without requiring action by the user.

In either case, the project should establish privacy policies and practices around collected attributes and be prepared to share those policies with institutions from whom they are requesting attribute release. (Ideally, they would publish those policies for their user communities as well.)

C.4.4 Level of Assurance

In Section A.4.2 we provided a brief introduction to the risks of federated identity; one of these risks is the risk that an identity provider, through some accident, misrepresents a user's identity. To help gauge this risk, InCommon has multiple levels of assurance that identity providers can operate under (e.g., Silver) [35]. An identity provider, by operating at a higher level of assurance, gives relying service providers greater assurance that identities are indeed consistent and reliable.

This document only covers the most basic of these levels and the more stringent levels are out of scope. One reason a project may need higher levels of assurance is if regulatory measures require it. The higher levels of assurance are meant to be consistent with the U.S. federal government's levels of assurance as defined in NIST Special Publication 800-63 [62]. NSF CI projects falling under the definition of Large Facilities should be aware of this decision to fulfill their requirements for a security plan on the terms of the cooperative agreement [64].

C.4.5 Sociological Impact of Outsourcing Authentication

A common challenge for existing projects adopting federated identity is that staff may resist the distribution of the identity management process across the InCommon federation. While there is no silver bullet to solving this problem, identifying key staff, providing them access to training and having them involving in the process as early as possible can mitigate this problem.

InCommon and NSF Cyberinfrastructure Glossary, References and Other Resources

D Glossary of Terms

For the reader's convenience, we provide here a set of terms relevant to federated identity used throughout this document. We thank the InCommon organization as many of these definitions are taken from the InCommon glossary [33] and reproduced with the permission of InCommon.

- **Administrator** - In the context of InCommon, the Administrator serves as the participating organization's primary registrar. The Administrator is responsible for registering and maintaining the policies and technical data related to the organization's participation in the InCommon Federation, including the submission of the URL of the Participant's POP and any Identity Provider and/or Service Provider metadata and associated certificates. The participating organization's designated Executive assigns the Administrator.
- **Assertion** - The identity information provided by an Identity Provider to a Service Provider.
- **Attribute** - A single piece of information associated with an electronic identity database record. Some attributes are general; others are personal. Some subset of all attributes defines a unique individual. Examples of an attribute are name, phone number, and group affiliation.
- **Attribute Assertion** - A mechanism for associating specific attributes with a user.
- **Attribute Authority (AA)** - The Shibboleth software service that asserts the requesting individual's attributes by creating an attribute assertion and then digitally signing it. The receiving online Service Provider must be able to validate this signature.
- **Attribute Release Policy (ARP)** - Rules that an AA follows when deciding whether or not to release an attribute and its value(s)
- **Audit** - An independent review and examination of a system's records and activities to determine the adequacy of system controls, ensure compliance with established security policy and procedures, detect breaches in security services, and recommend any changes that are indicated for countermeasures.
- **Authentication (AuthN)** - The security measure by which a person transmits and validates his or her association with an electronic identifier. An example of authentication is submitting a password that is associated with a user account name.
- **Authorization (AuthZ)** - The process for determining a specific person's eligibility to gain access to a resource or service, a right or permission granted to access an online system.

- **Boarding Process –** the term used to describe the process a service provider goes through on joining a federation to arrange receiving the attributes it requires from the identity providers.
- **Billing Contact** - In the context of InCommon, the Billing Contact is responsible for executing and maintaining all of the Participant's financial transactions associated with its InCommon federation participation, including any necessary communication with its internal Executive and Administrative Contacts, and externally with federation accounting staff.
- **Directory** - A directory is a specialized database that may contain information about an institution's membership, groups, roles, devices, systems, services, locations, and other resources.
- **eduPerson** - An LDAP object class authored and promoted by the EDUCAUSE/Internet2 eduPerson Task Force to facilitate the development of inter-institutional applications. The eduPerson object class focuses on the attributes of individuals. Current documentation on the eduPerson object class is available at http://www.educause.edu/eduperson/.
- **Electronic identifier** - A string of characters or structured data that may be used to reference an electronic identity. Examples include an email address, a user account name, a campus NetID, an employee or student ID, or a PKI certificate.
- **Electronic identity** - A set of information that is maintained about an individual, typically in campus electronic identity databases. May include roles and privileges as well as personal information. The information must be authoritative to the applications for which it will be used.
- **Enterprise directory** - An enterprise directory is a core middleware architecture that may provide common authentication, authorization, and attribute services to electronic services offered by an institution.
- **Executive –** In the context of InCommon, the Executive represents the participant organization regarding all decisions and delegations of authority for the responsibilities of InCommon Participants, including but not limited to payment of invoices, and assigning any person in the trusted Administrator role who submits Certificate Signing Requests, metadata, or Certificate Revocation Requests, and other administrative duties as described herein. The Executive is authorized as such in the InCommon participation agreement or by succession from the originally named Executive. The Executive role will typically be filled by a CIO, VP of IT, or other senior administrative officer responsible for the organization's information technology assets.
- **Federated identity** - The management of identity information between members of a federation.

- **Federation** - A federation is an association of organizations that come together to exchange information as appropriate about their users and resources in order to enable collaborations and transactions.
- **Identity** - Identity is the set of information associated with a specific physical person or other entity. Usually not all identity attributes are relevant in any given situation. Typically an Identity Provider will be authoritative for only a subset of a person's identity information.
- **Identity credential** - An electronic identifier and corresponding personal secret associated with an electronic identity. An identity credential typically is issued to the person who is the subject of the information to enable that person to gain access to applications or other resources that need to control such access.
- **Identity database** - A structured collection of information pertaining to a given individual. Sometimes referred to as an "enterprise directory." Typically includes name, address, email address, affiliation, and electronic identifier(s). Many technologies can be used to create an identity database or set of linked relational databases.
- **Identity Management System** - A set of standards, procedures and technologies that provide electronic credentials to individuals and maintain authoritative information about the holders of those credentials.
- **Identity Provider (IdP)** - The originating location for a user. Previously called the Origin Site in the Shibboleth software implementation. For InCommon, an IdP is a campus or other organization that manages and operates an identity management system and offers information about members of its community to other InCommon participants.
- **InCommon federation** - InCommon is a formal federation of organizations focused on creating a common framework for trust in support of research and education. The primary purpose of the InCommon federation is to facilitate collaboration through the sharing of protected network-accessible resources by means of an agreed-upon common trust fabric. InCommon participation is separate from membership in Internet2.
- **InCommon Technical Advisory Committee** - Group of individuals that provide technical guidance and direction for InCommon.
- **Metadata** - Data about data, or information known about an object in order to provide access to the object. Usually includes information about intellectual content, digital representation data, and security or rights management information.
- **Participant** - An organization accepted into the InCommon Federation that has met all the criteria for participation as either a higher education institution or a Sponsored Partner.

- **Participant Agreement (PA)** - This is the "contract" that a potential Participant signs when they are accepted by the Federation. This document outlines information such as fees, and responsibilities to participate in InCommon.
- **Participant Operating Practices (POP)** - This document describes how InCommon Participants need to describe their credential and identity management system.
- **Persistent Identifier** – A user identifier than is reused across multiple sessions. Such an identifier allows a service to maintain state about a user, for example, their ownership of data or personalization preferences.
- **Privacy Policy** - A statement to users of what information is collected and what will be done with the information after it has been collected.
- **Pseudonymous authentication** - Authentication with an identifier that remains consistent across sessions, but doesn't expose any personal information in itself, for example, a pseudonym one might create on an Internet forum.
- **Service Provider (SP)** - Previously called the Target Site in the Shibboleth software implementation. For InCommon, an SP is a campus or other organization that makes online resources available to users based in part on information about them that it receives from other InCommon participants.
- **Shibboleth®** - Software developed by Internet2 to enable the sharing of web resources that are subject to access controls such as user IDs and passwords. Shibboleth leverages institutional sign-on and directory systems to work among organizations by locally authenticating users and then passing information about them to the resource site to enable that site to make an informed authorization decision. The Shibboleth architecture protects privacy by letting institutions and individuals set policies that control what information about a user can be released to each destination. For more information on Shibboleth please visit: http://shibboleth.internet2.edu/uses.html.
- **Sponsored Partner** - A business partner that provides resources to a higher education institution, and is sponsored for participation in InCommon by a participating higher education institution.
- **Technical Contact** - The Technical Contact for InCommon serves as the primary point of contact for all technical issues for the organization participating in InCommon. The technical contact communicates with federation technical staff to ensure smooth operation of the federation's infrastructure.

E References

1. Accrediting Agencies Recognized by InCommon. http://www.incommonfederation.org/accrediting.html
2. Administrative Interface for Internet2 Services. https://service1.internet2.edu/siteadmin/manage/
3. Barton, Tom, Jim Basney, Tim Freeman, Tom Scavo, Frank Siebenlist, Von Welch, Rachana Ananthakrishnan, Bill Baker, Monte Goode and Kate Keahey. Identity Federation and Attribute- based Authorization through the Globus Toolkit, Shibboleth, GridShib, and MyProxy. 5th Annual PKI R&D Workshop, 2006.
4. Basney, Jim, Terry Fleury and Von Welch. Federated Login to TeraGrid. 9th Symposium on Identity and Trust on the Internet, 2010. http://www.ncsa.illinois.edu/~jbasney/tgfed.pdf
5. Central Authentication Service. Jasig. http://www.jasig.org/cas
6. CIC InCommon Silver Project: Phase 1 Report. July, 2010. http://www.cic.net/Libraries/Technology/IdM_InCommonSilverPhase1.sflb.ashx
7. CILogon Portal Delegation. http://www.cilogon.org/portal-delegation
8. CILogon Service. https://cilogon.org/
9. COmanage: Collaborative Organization Management. http://www.internet2.edu/comanage/
10. Configuring Shibboleth Delegation for a Portal. https://spaces.internet2.edu/x/n4Sg
11. Current InCommon Participants. http://www.incommonfederation.org/participants/
12. DataONE. https://www.dataone.org/
13. Developing a Coherent Cyberinfrastructure from Local Campus to National Facilities: Challenges and Strategies. A Workshop Report and Recommendations. EDUCAUSE Campus Cyberinfrastructure Working Group and Coalition for Academic Scientific Computation. February 2009. http://www.casc.org/papers/CASC-CCI_Workshop_Report_and_Recommendations.pdf
14. EDUCAUSE FedId Resources. http://www.educause.edu/Resources/Browse/FederatedIdentityManagement/31075
15. eduPerson and eduOrg Object Classes. http://middleware.internet2.edu/eduperson/
16. Family Educational Rights and Privacy Act (FERPA). http://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html
17. Goodin, Dan. Interpol chief impersonated on Facebook. The Register. September 20, 2010. http://www.theregister.co.uk/2010/09/20/interpol_chief_impersonated/
18. IdPAddAttribute. https://spaces.internet2.edu/display/SHIB2/IdPAddAttribute
19. IdPAddAttributeFilter. https://spaces.internet2.edu/display/SHIB2/IdPAddAttributeFilter
20. IdP Clustering Configuration. https://spaces.internet2.edu/display/SHIB2/IdPClusterIntro
21. IdPInstall. https://spaces.internet2.edu/display/SHIB2/IdPInstall
22. IdPLogging. https://spaces.internet2.edu/display/SHIB2/IdPLogging
23. IdPMetadataProvider. https://spaces.internet2.edu/display/SHIB2/IdPMetadataProvider
24. IdPUserAuthn. https://spaces.internet2.edu/display/SHIB2/IdPUserAuthn
25. InCommon: About InCommon. http://www.incommon.org/about.html
26. InCommon Affiliates. http://www.incommonfederation.org/affiliate/

27. InCommon Basics and Participating in InCommon: A Summary of Resources. https://spaces.internet2.edu/download/attachments/2815/resources_booklet.pdf?version=3
28. InCommon Benefits. http://www.incommonfederation.org/benefits.cfm
29. InCommon Cert Service. http://www.incommonfederation.org/cert/
30. InCommon Education and Training. http://www.incommonfederation.org/educate/
31. InCommon Fee Structure. http://www.incommonfederation.org/fees.html
32. InCommon for Service Providers. http://www.incommon.org/partners/
33. InCommon Glossary. http://www.incommonfederation.org/glossary.cfm
34. InCommon IAM Online Presentations http://www.incommonfederation.org/iamonline/
35. InCommon Identity Assurance. http://www.incommonfederation.org/assurance/
36. InCommon Introductory Presentation https://spaces.internet2.edu/download/attachments/2815/InC_Overview_v2.ppt?version=1
37. InCommon Library Collaboration. http://www.incommonfederation.org/library/
38. InCommon Metadata Consumption. https://spaces.internet2.edu/display/InCCollaborate/Metadata+Consumption
39. InCommon Policies and Practices. http://www.incommonfederation.org/policies.html
40. InCommon Sponsored Partners. http://www.incommonfederation.org/sponsor.html
41. InCommon Shibboleth Metadata Configuration. https://spaces.internet2.edu/display/InCCollaborate/Shibboleth+Metadata+Config
42. InCommon Technical Guide https://spaces.internet2.edu/display/InCCollaborate/Technical+Guide
43. Indiana Clinical and Translational Sciences Institute (CTSI). http://www.indianactsi.org/
44. Information from InCommon. https://spaces.internet2.edu/display/InCCollaborate/Information+from+InCommon
45. Install Shibboleth to protect Java Servlets. https://spaces.internet2.edu/display/SHIB2/NativeSPJavaInstall
46. Internet2: Boarding Process. https://spaces.internet2.edu/display/fedapp/Boarding+Process
47. Internet2 Events. http://events.internet2.edu/
48. Internet2: FedApps Working Group. https://spaces.internet2.edu/display/fedapp/Home
49. Internet2: NIH Federation InCommon Wiki. https://spaces.internet2.edu/display/InCNIH/InC-NIH
50. Inernet2: OpenID Use Cases. https://spaces.internet2.edu/display/OpenID/Use+Cases
51. JISC Idm Toolkit. https://gabriel.lse.ac.uk/twiki/bin/view/Projects/IdMToolkit/Toolkit
52. Johnson, H. and P. Caskey. Introduction to Shibboleth Attribute Delivery. Educause CAMP, June 2007. http://www.educause.edu/Resources/IntroductiontoShibbolethAttrib/161780
53. Join InCommon. http://www.incommonfederation.org/join.cfm
54. Laser Interferometer Gavitational-Wave Observatory (LIGO). http://www.ligo.caltech.edu/
55. Leve, Kristina and Valter Nordth. Lowering costs of identity proofing by federated identity management. http://www.incommonfederation.org/docs/other/SWAMI_federated_idm_roi.pdf
56. Metadata. https://spaces.internet2.edu/display/SHIB2/Metadata
57. Morgan, R. L. "Bob", Scott Cantor, Steven Carmody, Walter Hoehn, and Ken Klingenstein. “Federated Security: The Shibboleth Approach.” Educause Quarterly, Volume 27,

Number 4, 2004. http://www.educause.edu/EDUCAUSE+Quarterly/EDUCAUSEQuarterlyMagazineVolum/FederatedSecurityTheShibboleth/157315

58. National Science Foundation and Penn State InCommon Pilot Now Underway. November 17, 2010. https://iam.psu.edu/national-science-foundation-and-penn-state-incommon-pilot-now-underway
59. National Science Foundation Cyberinfrastructure Council. Cyberinfrastructure Vision for 21st Century Discovery. March 2007. http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf
60. NativeSPClustering. https://spaces.internet2.edu/display/SHIB2/NativeSPClustering
61. NativeSPMetadataProvider. https://spaces.internet2.edu/display/SHIB2/NativeSPMetadataProvider
62. NIST Special Publication 800-63: Electronic Authentication Guideline. Version 1.0.2. April, 2006.
63. NITLE Shib and FedId Roadmap for Smaller Colleges and Universities. http://cnx.org/content/m31491/latest/
64. NSF Cooperative Agreement: Supplemental Financial/Administrative Terms and Conditions--Large Facilities. September 25, 2006
65. NSF Dear Colleague Letter: Cyberinfrastructure Framework for 21st Center Science and Engineering (CF21). http://www.nsf.gov/pubs/2010/nsf10015/nsf10015.pdf
66. NSF Resource Allocations Policies. https://www.teragrid.org/web/user-support/allocations_policy
67. OASIS Security Services (SAML) Technical Committee. http://www.oasis-open.org/committees/security/
68. OAuth: Introduction. http://oauth.net/about/
69. Ocean Observatory Initiative (OOI). http://www.oceanleadership.org/programs-and-partnerships/ocean-observing/ooi/
70. Open Science Grid. http://www.opensciencegrid.org/
71. OSG Registration Authority. https://twiki.grid.iu.edu/twiki/bin/viewauth/OSGRA/
72. Project Moonshot. http://www.project-moonshot.org/
73. ProtectNetwork. http://www.protectnetwork.org/
74. REFEDS: Research and Education Federations. http://www.terena.org/activities/refeds/
75. Science Gateways Home. https://www.teragrid.org/web/science-gateways/
76. Shib Enabled. Internet2. https://spaces.internet2.edu/pages/viewpage.action?pageId=11484
77. Shibboleth 2 Identity Provider Configuration. Internet2. https://spaces.internet2.edu/display/SHIB2/IdPConfiguration
78. Shibboleth: About http://shibboleth.internet2.edu/about.html
79. Shibboleth: Communicating with a Service Provider. https://spaces.internet2.edu/display/SHIB2/IdPSPCommunicate
80. Shibboleth Deployment Checklist http://shibboleth.internet2.edu/shib-checklist-final-website.html
81. Shibboleth Getting Started http://shibboleth.internet2.edu/get-started.html
82. Shibboleth Installation. https://spaces.internet2.edu/display/SHIB2/Installation
83. Shibboleth Mailing Lists. http://shibboleth.internet2.edu/lists.html
84. Shibboleth Support Documentation http://shibboleth.internet2.edu/support.html
85. Shibboleth: Talk to a New IdP. https://spaces.internet2.edu/display/SHIB2/NativeSPAddIdP

86. Shibboleth Technical Deployers Info Center http://shibboleth.internet2.edu/deployers.html
87. Shibboleth Technical Manager Info Center http://shibboleth.internet2.edu/adopters.html
88. Siebenlist, F., R. Ananthakrishnan, D. E. Bernholdt, L. Cinquini, I. T. Foster, D. E. Middleton, N. Miller, and D. N. Williams, "Enhancing the Earth System Grid Security Infrastructure through Single Sign-On and Autoprovisioning," Proceedings of the 5th Grid Computing Environments Workshop, Portland, Oregon, USA, ACM, 2009. http://www.mcs.anl.gov/uploads/cels/papers/P1683.pdf.
89. Support. http://shibboleth.internet2.edu/support.html
90. SWITCH AAI Demo. http://www.switch.ch/aai/demo/
91. TeraGrid: Campus Champions. https://www.teragrid.org/web/eot/campus_champions
92. TeraGrid User Portal. https://portal.teragrid.org/
93. uApprove. SWITCH. http://www.switch.ch/aai/support/tools/uApprove.html
94. Welch, Von, Tom Barton, Kate Keahey and Frank Siebenlist. Attributes, Anonymity, and Access: Shibboleth and Globus Integration to Facilitate Grid Collaboration. 4th Annual PKI R&D Workshop, 2005
95. Welch, V., Frank Siebenlist, Ian Foster, John Bresnahan, Karl Czajkowski, Jarek Gawor, Carl Kesselman, Sam Meder, Laura Pearlman and Steven Tuecke. Security for Grid Services. Twelfth International Symposium on High Performance Distributed Computing (HPDC-12), 2003.
96. What is OpenID? http://openid.net/get-an-openid/what-is-openid/
97. X.509 Certificates in Metadata. Internet2. https://spaces.internet2.edu/display/InCCollaborate/X.509+Certificates+in+Metadata

F Additional Resources

F.1 Future Resources

In this section we briefly discuss technologies and other services that are not available today, but are expected to be available in the next future and which may bring benefit to NSF CI projects and institutions that house NSF researchers. These technologies are listed in no particular order.

F.1.1 UApprove

uApprove [93] is a software extension to a Shibboleth IdP which allows the user to make attribute release decisions instead of relying on the Shibboleth IdP administrator and organization policy. Many organizations are hoping that by putting the decision into the hands of the users, it will ease concerns around FERPA [16].

F.1.2 Federated SSH Work

A current limitation of InCommon and Shibboleth is a lack of support for applications other than web browsers. As we discussed in the Guide to Technical Deployment, CILogon exists to bridge from InCommon to PKI credentials, used by many command line grid applications.

Two future developments that are working to address adding federating identity support to a broader range of non-web application are Project MoonShot [72] and the Federated SSH work as part of the COmanage project [9].

F.1.3 FedApps Working Group

Internet2 is starting a working group to investigate issues involved with making applications available via federated identity. This working group, entitled “FedApps” [48], is in the process of forming at the time of this writing.

F.2 Identity Management Resources

Establishing an identity management system is outside the scope of this document, some resources for doing are:

- The NMI-Edit web site: http://www.nmi-edit.org/started/index.cfm
- InCommon IAM Online: http://www.incommon.org/iamonline/
- Educause Federated Identity Management Resources: http://www.educause.edu/Resources/Browse/FederatedIdentityManagement/31075
- Jansson, Eric. NITLE Shibboleth and Federated Identity Management Roadmap for Smaller Colleges and Universities. Connexions. August 20, 2009. http://cnx.org/content/m31491/latest/

- JISC. The Identity Management Toolkit. https://gabriel.lse.ac.uk/twiki/bin/view/Projects/IdMToolkit/Toolkit

F.3 Resources for Federated Identity Deployment

F.3.1 Examples of Deployments

In this section we list some examples of deployments to provide the reader with some real-world experience from institutions that may approximate their own:

- EDUCAUSE presentation describing experiences at UCLA, Penn State and New Castle (UK): http://www.educause.edu/Resources/ShibbolethCaseStudies/161773
- Deploying Shibboleth: Technical Requirements, Policy Issues, and Case Studies (presentations from USC, Penn State, MIT): http://www.educause.edu/Resources/DeployingShibbolethTechnicalRe/169205
- Shibboleth In Use, a collection of use cases on the Shibboleth web site: http://shibboleth.internet2.edu/shib-in-use.html
- InCommon Case Studies: http://www.incommonfederation.org/cases.html
- USC Case Study by NMI-EDIT: http://www.nmi-edit.org/case_studies/usc-shibpubc.pdf
- Implementing a production Shibboleth IdP service at Cardiff University (presentation slides): http://www.slideshare.net/JISC.AM/cardiff-jisc-fam-aston-may07
- InCommon … Now That's the Ticket. Lafayette provides students with SSO ticketing convenience. http://www.incommonfederation.org/docs/eg/InC_CaseStudy_UTix_Lafayette_2009.pdf

F.3.2 InCommon Training Opportunities

The best place to look for an up to date list of training opportunities is the InCommon Education and Training web site [30], which includes both a list of in-person workshops and online seminars.