Federated Security: The Shibboleth Approach

The open-source Shibboleth System extends Web-based applications and identity management for secure access to resources among multiple organizations

By **R. L. "Bob" Morgan, Scott Cantor, Steven Carmody, Walter Hoehn,** and **Ken Klingenstein**

The Fifth Annual Educause Current Issues Survey[1] ranked "security and identity management" near the top of the list of critical IT challenges on campus today. Recognition of the crucial importance of securing networked resources led Internet2 to establish its Middleware Initiative (I2MI) in 1999. While Internet2 was founded to develop and deploy advanced network technologies and applications, it was clear from the start that high-speed networks would simply provide a quicker path to abuse unless improved methods of managing and controlling access to resources were developed and deployed along with those networks. I2MI has brought together campus middleware architects to work on fundamental issues in authentication, authorization, and directory services to make secure inter-institutional services possible and practical.

The most innovative I2MI effort to date is the Shibboleth Project.[2] Its primary product, the Shibboleth System,

- supports secure user access to Web-based resources;
- enables independent organizations to federate to extend the capabilities of their existing identity-management services;
- supports multi-organizational federations to enable scalable use of the technology;
- encourages attribute-based authorization;
- provides controls to protect the privacy of personal information;
- is standards-based and open-source;
- has entered production use; and
- is evolving to support new uses and new communities.

Overview

The Shibboleth System is often called just Shibboleth. Here we describe how it works, its key features, and how it is designed to meet the needs of the higher education and research communities and their partners.

How It Works

The Shibboleth System includes two major software components: the Shibboleth Identity Provider (IdP) and the Shibboleth Service Provider (SP). These two components are deployed separately but work together to provide secure access to Web-based resources.

A step-by-step description of the Shibboleth sign-on process follows. While the details may vary based on deployment choices, the steps below are typical. The players include the user, who wants to use a protected Web resource; the resource provider Web site, which has installed the Shibboleth SP software; and the user's home organization, which has installed the Shibboleth IdP software.

1. The user navigates to the Web resource using her browser. The resource site is protected, hence requires information about the user in order to decide whether access is permitted.

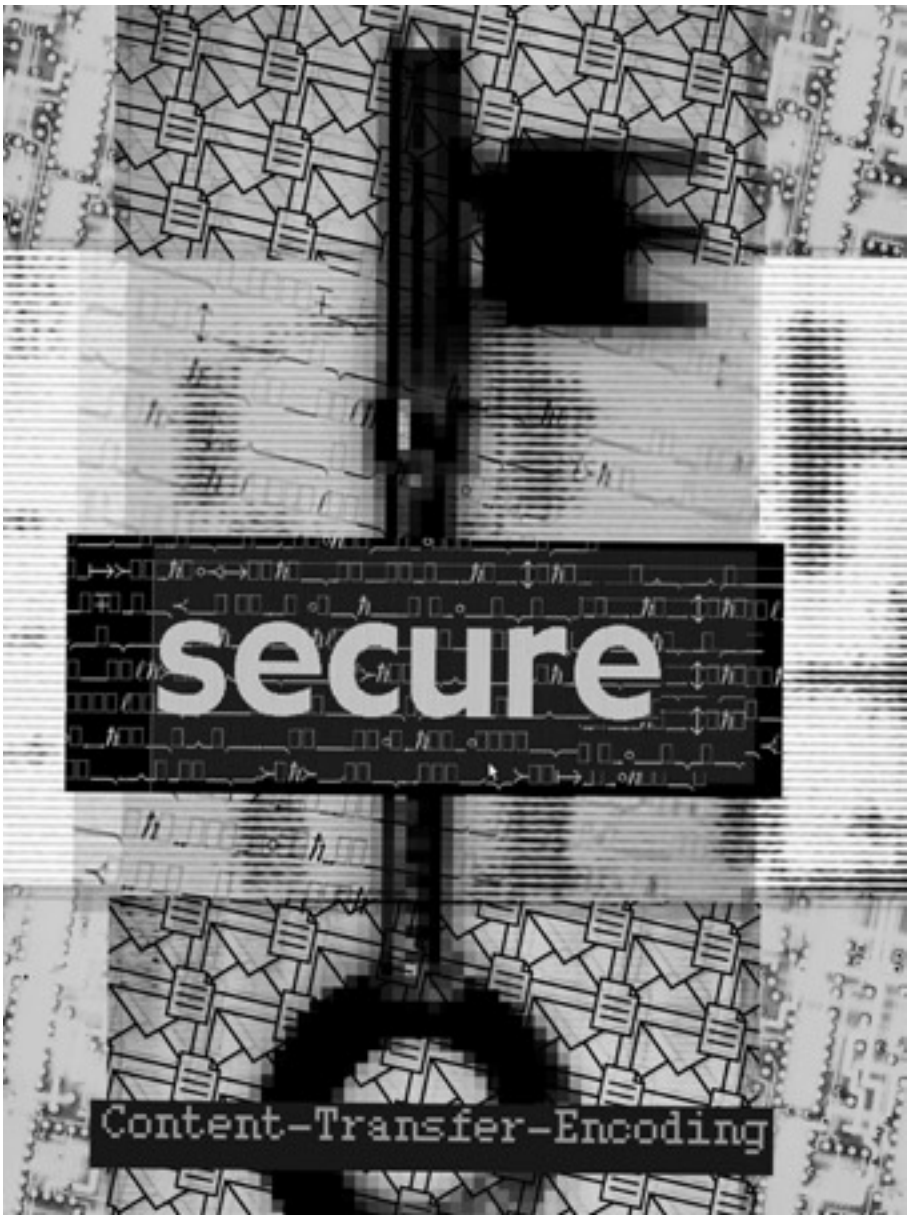
2. The Shibboleth SP software redirects the browser to a "navigation" page (called a WAYF, for "where are you from"), which presents the user with a list of the organizations whose users may access the resource.

3. The user selects her home organization, and the browser is sent to the home organization's Web site running the Shibboleth IdP software. This site uses a Web sign-on method chosen by the home organization. The user now sees the familiar login Web page of her home organization, enters her username and password, and selects the Login button.

4. The Shibboleth IdP software sends the browser back to the original resource site and includes in the message some security information called an "assertion" that proves the user signed on. The Shibboleth SP software on the resource site validates the assertion and then requests additional information (attributes, such as "faculty" or "student in Film327") about the user by making a request to the home organization's Shibboleth IdP service.

5. The Shibboleth SP receives the user's attributes from the home organization's IdP and passes them along to the resource provider's Web application. The application uses those attributes and its access policy to decide whether the user's access is permitted or denied, displaying the appropriate page to the user's browser.

Often, many of these steps can be skipped. The WAYF can set a cookie in the user's browser so that the user doesn't see that page the next time through. If the home organization's Web authentication service uses single



sign-on and the user already has a session with it, the login page won't be seen. In many cases the user can get access to the resource without seeing any intermediate Web pages at all.

The process above resembles other Web sign-on schemes. In the rest of this section we present the features that distinguish the Shibboleth System.

The SAML Standard and Federated Identity

The operation of the Shibboleth System is based on, and conforms to, the Security Assertion Markup Language (SAML) standard (version 1.1), published by OASIS (http://www.oasis-open.org/), the leading standards body for technology based on the Extensible Markup Language (XML). SAML was created by many leading security experts from industry and academia, including members of the Shibboleth Project, specifically to provide interoperability among Web sign-on products, many of which now support SAML. Using SAML permits Shibboleth to work with many vendor products and gives it a solid technical foundation as the standard evolves.

The principle behind SAML's design—and Shibboleth's—is federated identity. One of the Internet's key strengths is its media independence—the ability of an Internet Protocol (IP) packet can to travel across many different physical networks. Similarly, federated identity technology allows organizations using disparate authentication and authorization methods to interoperate, extending the capability of each organization's existing services rather than forcing their replacement. Federated identity also helps users by taking advantage of their familiarity with existing sign-on systems and reducing the number of passwords users have to remember.

Attribute-Based Authorization

Typical user authentication methods only provide the application with the permanent user identifier (userid) of the person who has authenticated. This simple approach won't suffice in modern systems. Applications need additional information about users—user attributes—to make proper authorization decisions. Providing this information as part of the sign-on process is especially useful in multi-organizational situations where an application probably won't have access to user information through other means such as a directory service. Shibboleth is designed specifically to provide user attributes to applications with the flexibility, extensibility, security, and privacy required in federated scenarios. Organizations can use Shibboleth's built-in attribute support (based on the Internet2/EDUCAUSE eduPerson directory schema, <http://www.educause.edu/eduperson/>) or create new attributes to meet the needs of applications. For example, attributes can represent "entitlements" such as "user is authorized to access resource collection X."

The Shibboleth IdP software plugs in to existing institutional identity management and user information services (typically Lightweight Directory Application Protocol, or LDAP-based, directories), extending them to work inter-organizationally.

User Privacy Protection

A key difference between intra-organizational and multi-organizational systems is the strong requirement for protection of personal information in the latter, as reflected in federal legis-

lation and university privacy policies. These principles guided the Shibboleth design: the users should control what personal information is released and to whom, and the resource provider should only receive as much user information as needed to make access control decisions unless the user chooses to release more. It is also very important that privacy protection not hinder use—it must be as easy to release information as to protect it.

Shibboleth's emphasis on user attributes is an important tool in privacy protection. In many systems a user login gives the resource provider a well-known "userid" (often also used as an e-mail address), making privacy leakage inevitable. In Shibboleth the userid is just another attribute, only sent if access to the resource requires it. If, as in many scenarios, only a membership attribute is needed, that's all the resource provider will get.

The Shibboleth IdP software has a key subsystem for management of attribute release policies, permitting fine-grained control of information release based on which resource provider is receiving the attributes. Since managing these policies will likely be a burden for the average user, Shibboleth provides methods for establishing defaults and administrator control. The Shibboleth Project is actively researching ways of making privacy management easy for the typical user.

Federations

Given the flexible nature of modern standards and software, those deploying the Shibboleth System must make many choices. For organizations to successfully interoperate using federated identity, for example, they must agree on many technical points, including the following:

- security mechanisms used among the Shibboleth servers (usually X.509-based Public Key Infrastructure, PKI),
- definition of attributes, and
- how to locate the servers of other participants.

They must also agree on higher-level policy questions, such as

Shibboleth's emphasis on user attributes is an important tool in privacy protection.

- the accuracy of user-management practices,
- procedures for handling sensitive personal information, and
- the sorts of organizations that may participate.

Clearly, making these agreements once to meet many needs of a large community scales much better than relying on a myriad of two-party arrangements. In the Shibboleth Project a community based on such agreements is called a federation.

The Shibboleth System supports federations by defining formats for managing site configuration information and providing procedures for creating, distributing, and importing that information. Beyond this, a Shibboleth SP or IdP might need to participate in multiple federations as well as two-party arrangements, so the Shibboleth software permits sophisticated configurations where a service may have many policies active simultaneously.

As part of its overall mission of meeting the needs of its higher-education members, Internet2 has established the InCommon Federation (http://www.incommonfederation.org/) as a formal federation of organizations focused on creating a common framework for trust in support of research and education. InCommon supports the use of Shibboleth software by its participants, both identity providers (primarily U.S. higher-education sites) and resource providers (partners such as commercial information and service providers, as well as higher-education resource sites).

Federations in other communities are already in progress, including research- and education-based federations in countries such as Finland, Switzerland, and the United Kingdom. Federations in other communities, such as U.S. state governments, industry consortia, and other partnerships, are under active discussion.

The Shibboleth Project

Like all I2MI projects, the Shibboleth Project uses open design and development processes. The project has benefited from contributions by dozens of participants from academic organizations around the world and from partners in industry. In particular, the digital library community has provided significant help clarifying licensed-content scenarios and generating interest from commercial content services. The Shibboleth System is open source software, using nonrestrictive licensing terms to promote its wide adoption in open-source and proprietary products.

With a growing number of partners and increasing adoption in the international community, the Shibboleth Project has recognized the need to broaden involvement in its governance process, specifically to coordinate significant new investments in the system. (See the sidebar for status of the Shibboleth System's deployment.)

Shibboleth in Action

Many organizations are using the Shibboleth System today to solve multi-organizational Web access problems. Many other applications are in the pilot stage. Some examples follow.

A Student-Oriented Information Service

Pennsylvania State University's arrangement to provide students with access to the Napster music service has been covered extensively in the press. The Shibboleth System played a key role in making this service accessible while meeting both the university's and Napster's security and privacy requirements.

Like many universities, Penn State has an existing infrastructure supporting sign-on to Web-based applications using a campus network identifier (userid) system. Using this for sign-on to the Napster service was unappealing for various reasons. First, the scheme was specific to Penn State and so wouldn't apply at other universities with which

Shibboleth Deployment Status

Version 1.0 of the Shibboleth System was released in June 2003. As of this writing the current version, 1.2, is in use or in test by more than 150 organizations, including universities, research labs, commercial service providers, and software vendors. The first wave of licensed content providers have begun supporting Shibboleth-enabled access, including JSTOR (http://www.jstor.org/), OCLC (http://www.oclc.org/), EBSCO (http://ebsco.com/), and Elsevier's ScienceDirect service (http://www.sciencedirect.com/). The second wave of implementations is currently under way. Discussions have begun with vendors offering outsourced Web-based services in several different application spaces.

Internationally, Shibboleth is deployed throughout Switzerland by SWITCH (the Swiss Education and Research network, <http://www.switch.ch/>) and in Finland. Recently the United Kingdom's Joint Information Systems Committee (JISC, <http://www.jisc.ac.uk/>) funded eight projects related to Shibboleth deployment across higher education, along with further Shibboleth software development. The Australian higher education community is currently pursuing country-wide deployment.

Several non–Web-based projects—such as instant messaging, peer-to-peer resource sharing, and grid systems—are actively exploring Shibboleth integration. A joint effort with Microsoft is under way to provide interoperability with the IBM-Microsoft Web Services Security Model. Finally, Shibboleth is in the process of being certified for use with the U.S. Federal E-Authentication Initiative (http://cio.gov/eauthentication/).

Napster hopes to make similar arrangements. Second, a userid-based sign-on would expose those userids to Napster, an undesirable practice for any externally delivered service without a strong requirement for them. Third, only a subset of students were permitted to use the service, so Penn State would somehow have to tell Napster which students were authorized without revealing their userids.

The Shibboleth System addressed all these concerns. Napster deployed the standard Shibboleth SP software. Penn State had an existing Shibboleth IdP service, which it expanded to meet expected load from this application. The university used a Shibboleth feature called the "targeted ID" attribute to provide Napster with a persistent user identifier at login that was unrelated to the student's userid or e-mail address. Penn State also provided an "OK for Napster" attribute for authorized users, so Napster could recognize them dynamically. The service successfully went live early in 2004.

An Academic Information Provider

JSTOR (http://www.jstor.org/) is a nonprofit organization with a dual mission: to create and maintain a trusted archive of important scholarly journals, and to provide access to these journals as widely as possible. It provides controlled access to its archive to researchers, librarians, faculty, and staff at participating institutions. Like many similar information services, JSTOR does its access control primarily using network addresses. A participating institution tells JSTOR its network address ranges, and JSTOR's servers permit access by all computers at those addresses.

A number of well-known problems afflict this access-control method, but when implemented, it was the only practical choice. Among the most dangerous problems is that if one of the computers on a campus is compromised, that machine can be used for unauthorized access to JSTOR by a remote exploiter. This exploit has happened, causing JSTOR to expend significant resources detecting and defending against it.

JSTOR was an early and enthusiastic Shibboleth Project participant, seeing Shibboleth as the best approach for moving beyond IP address-based access control. By requiring an interactive login for access, Shibboleth would prevent the exploit described above. In addition to better security, access via Shibboleth would provide JSTOR with better opportunities for personalized services, without requiring JSTOR-specific accounts and passwords. JSTOR is also interested in applying Shibboleth and SAML technology to solve problems with proxy-style search access to repositories.

JSTOR has worked with several campuses on a trial of Shibboleth-enabled access to its archives. It is anticipating going live as part of the InCommon Federation.

A Research Collaboration

Academic research projects, especially in the sciences, increasingly involve sophisticated computing resources and participants from multiple institutions. These virtual organizations (VOs), set up to support scientists doing research, find themselves confronting the same identity-management issues as campus IT shops—password distribution and resetting, levels of assurance, authorization management—but with users spread across the country or around the globe and with limited staff to do the work. It's no wonder that attention is now going to methods that will help VOs by letting them rely on existing campus IT services. By using Shibboleth to control access to its Web-based resources, a VO can accommodate users (and useful attributes such as "faculty") coming from campuses with Shibboleth-enabled identity services.

In practice, many scientific VOs use both Web and non-Web technologies for resource access. Grid technology (http:

//www.ggf.org/) is an important platform for scientific computing, with its own approach to the multi-institutional access problem. Grid security researchers and Shibboleth Project members are working together on integrating the two infrastructures. This work focuses on having the Shibboleth IdP service provide user attributes to non-Web–based grid applications. The modularity of the Shibboleth System makes this possible. We expect the results of this work to become available in 2005.

An Outsourced Employee Application

As campus IT organizations consider how to deploy new administrative applications cheaply and flexibly, they find hosted services to be more common and more compelling. In such areas as procurement, charitable giving, and benefits management, vendors offer Web-based access to services used by potentially thousands of campus staff. Managing sign-on to these services is a key consideration in whether the hosted approach succeeds. Adding yet another username and password is unappealing, as is having remote services handle campus userid passwords. Vendors these days understand the appeal of single sign-on to campuses but typically invent their own sign-on schemes—of questionable security and uncertain supportability.

Shibboleth provides a solution for these applications also. Vendors that serve higher education now find themselves asked to support a different Web sign-on system by each campus, increasing the cost and complexity of their services. Some vendors turn to commercial Web sign-on products, but campuses are often reluctant to license commercial software for this purpose. Shibboleth represents a common, well-supported method for enabling what many expect to be an increasingly popular arrangement. In addition, Shibboleth's use of attributes permits campuses to express roles like "purchasing agent" or "benefits-eligible retiree" to providers at sign-on, eliminating delay-prone batch-feed methods of maintaining this information.

Managing sign-on to these services is a key consideration in whether the hosted approach succeeds.

Extended User Populations

Many campuses face the problem of providing secure access to an ever-expanding set of resources to new user populations beyond the traditional groups—students, faculty, and staff. Alumni services is the most common case, but other groups include retirees, students at other local colleges, K–12 students, contractors, extended education students, and medical providers. Campus-run applications such as portals, publishers, outreach programs, medical-information services, and others want to provide controlled access to these groups. In some cases campuses just treat these users as yet more participants in their regular user identity space, but this can create serious strain on policies and procedures for core identity management services.

Many campuses are looking at Shibboleth to help solve this problem. A Shibboleth-enabled identity provider can be set up for the extended population (or one for each of several), which can then have its own policies, procedures, and brand without affecting core identity services. Web-based applications that need to serve the extended population can use the Shibboleth SP software and accept users from both the regular campus and extended IdPs.

As an example, Columbia University Digital Knowledge Ventures (DKV) develops and distributes digital resources beyond the Columbia campus. Via their Columbia Educational Resources Online (CERO) Web site, content is available to the Columbia community for free and to individual and institutional subscribers for a fee. DKV recast CERO's customer database as a Shibboleth IdP and enabled the CERO site with Shibboleth SP software. Now the site can be accessed by either community using a common scheme.

As another example, the University of Washington's Catalyst system (http://catalyst.washington.edu/) offers educators an integrated collection of tools to make effective use of Web-based instruction. A community college in the Seattle area is interested in using Catalyst for its courses. By adding the Shibboleth SP software to the Catalyst site and installing the Shibboleth IdP software at the community college, the college's students can sign on to Catalyst with their community college userids. This approach is much easier than replicating the Catalyst system or providing the students with UW userids. It is also more scalable as other local institutions use Catalyst in the future.

Conclusions

Shibboleth provides an effective solution for secure multi-organizational access to Web resources. IT managers might think that this technology sounds complicated and that external access isn't at the top of the priority list right now. We suggest that the implications of Shibboleth and its adoption by many campuses and service providers has compelling implications even for those not on the bleeding edge:

- *Meeting campus identity management standards*: Federations such as InCommon are establishing a baseline for campus infrastructures to participate in multi-institutional scenarios. This is strong motivation for IT organizations to bring their local services up to standard in the areas of campus-wide user authentication and directory data management. In particular, stating assurance levels regarding how users authenticate (for example, strength of passwords) is important to applications both external and internal.
- *Privacy control*: Shibboleth meets strong requirements from higher-education communities to provide appropriate protection of personal information even when services use access control. Dealing effectively with privacy concerns is particularly complex; technology is only part of the story. Shibboleth provides campuses with controls so that as a community we can determine the bal-

ance points. By deploying Shibboleth, campuses can take better part in this discussion.

- *Attribute-based authorization*: Managing authorization in a secure and cost-effective fashion is a major goal for many IT organizations. The handling of attributes in Shibboleth provides a testbed for the use of role- and attribute-based authorization for applications of all kinds, not just multi-organizational ones. IT organizations can benefit from learning about and contributing to this emerging practice.

A large and growing community is using the Shibboleth System to solve problems and enable a new generation of applications and services. We encourage organizations of all kinds to try out the Shibboleth System and participate in the Shibboleth Project. *e*

Endnotes

1. D. Z. Spicer, P. B. DeBlois, and the EDUCAUSE Current Issues Committee, "Fifth Annual EDUCAUSE Survey Identifies Current IT Issues," *EDUCAUSE Quarterly,* Vol. 27, No. 2, 2004, pp. 8–22, <http://www.educause.edu/eq/eqm04/eqm0422.asp>.
2. For more information about the Shibboleth System and the Shibboleth Project and to download the software, see <http://shibboleth.internet2.edu/>.

R. L. "Bob" Morgan (rlmorgan@washington.edu) is Senior Technology Architect in the Computing and Communications division at the University of Washington in Seattle and is chair of the Internet2 Middleware Architecture Committee for Education (MACE). Scott Cantor is a Senior Systems Developer in the Office of Information Technology at The Ohio State University in Columbus, Ohio. Steven Carmody is an IT Architect in Computing and Information Services at Brown University in Providence, Rhode Island. Walter Hoehn is an Application Developer for the Information Technology Division at the University of Memphis in Tennessee. Ken Klingenstein is Chief Technologist at the University of Colorado in Boulder and Project Director for the Internet2 Middleware Initiative.