PrintFullVersion

Note: Please ignore the "date last changed" at the top of this page; that indicates when this page was last edited, not the last time that an included page was edited. this page INCLUDEs all of the active pages in the wiki, and allows you to create a printed version.

tom scavo has an include -- DRAFT

Unable to render {include} The included page could not be found.

Federated Enterprise Identities

Campuses have been assigning Digital Identities and associated Credentials to members of their communities for 10-20 years. These usually take the form of a userid and password. Initially, the identities were used primarily to access email. More recently, they are being used to access a wide range of both in-sourced and out-sourced business systems and systems supporting instruction. Students use them to submit course work; instructors use them to review submitted work and to post grades. In the last few years, a growing number of campuses have deployed Web Single SignOn system that provide an "authenticate once" function when accessing this broad array of web systems. Federated Business Identities (as asserted by both campuses and businesses) have been using the OASIS SAML protocol to carry Assertions.

The vast majority of campuses have created Business Processes to ensure that a Digital Identity is given to the person whose real world identity is associated with that Digital identity. Similar processes have existed for years with payroll systems (eg the I-9 process). Similarly, campuses have created Business Processes requiring a person to "prove" their identity as part of obtaining a Digital Identity. The goal is to ensure that a person submitting course work as John Smith is indeed John Smith. And that an Instructor submitting grades as Jane Doe is indeed Jane Doe. Consequently, these accounts are viewed as having a sufficiently high Level of Assurance to allow them to be used for business transactions within the enterprise that issued accounts. Increasingly, these accounts are being used to authenticate to gain access to information and services at business partners (agents) of the campus.

In recent years, however, campuses have also been issuing credentials to communities and people outside their core communities. These communities may have a weaker relationship to the campus (eg applicants, parents); individual people may be partnering with researchers on the campus. Many of the people in these new groups, however, are physically remote from the campus, and consequently the process of linking a person to these accounts is probably much weaker than the process used for members of the core communities. Interestingly, many of the people in these new communities already have identities issued by social identity providers; however, campuses have only begun to look at leveraging these social identities rather than issuing new identities that have an LoA at about the same level as the social providers. Consequently, Service Providers can no longer rely on a successful campus authentication (with no other information) as meaning that the user has a strong association with the campus. The current reality is that campuses can authenticate a large pool of people that includes their core communities and many others, and that these authentications are done at a broad spectrum of Assurance Levels.

Because of their Registration processes and Business processes, campuses can, however, associate meaningful information with each user. Beyond merely asserting that a successful authentication has taken place, campuses can differentiate members of the core community from these other groups by asserting attributes that describe the person's relationship with the campus; Social Identity Providers cannot currently do that. Minimally, a campus can assert whether a person is a faculty member, student, or staff, or a member of one of these other populations. If the person is a member of the core community, the campus can reliably assert additional information such as name and permanent identifiers.

Currently, campus identity providers and the social identity providers are categorized as operating at "LoA 1" (see TN -- Levels of Assurance). Currently, no US campuses are certified as operating at LoA 2. However, because the campuses have business processes in place as part of their Registration and Account Issuance Business Processes that include identity vetting, many people feel that campus issued identities are "stronger" than social identities (which do not have comparable processes). The Australian Federation considers campus-issed credentials to be at "LoA 1.5" (which is not defined anywhere...).

With userids and passwords, users can decide to share these values with other people, thus allowing other people to access systems using their Digital Identity. Multiple people can know the password, and can use it simultaneously. This is a problem with any Identity service that relies on passwords. Consequently, the above sentence has to be tempered to "an Instructor submitting grades as Jane Doe is indeed Jane Doe (or someone who knows Jane's password, either because Jane shared the password or they cracked Jane's weak password)".

As a result, Service Providers have adopted a "Risk Assessment" framework. See TN – Levels of Assurance for a detailed description. Basically, an SP performs an evaluation to determine the severity of an incident where an imposter is able to login to their site. The result of that analysis allows them to specify a required "Level of Assurance" for the Credentials that are used to login to the site. The campus can use Profiles to evaluate their Policy, Business Practice, and technology for a set of Credentials to determine their associated LoA. A user can access a specific SP if they are able to authenticate using Credentials with the required LoA. As a result, some campuses now have some set of people able to authenticate with mechanisms that are thought to be "stronger" than just userid/passwords.

Social Identities

In the past, applications owners would add a "new user registration" process to their site, and would issue userids and passwords to these "outsiders". This created a burden for both sides -- the user would have to remember yet another set of credentials, and the site would have to institute business processes to deal with forgotten passwords, etc.

A growing number of these applications, though, are looking to "outsource" the identity problem by leveraging the authentication and Web Single SignOn (SSO) functionality provided by the big internet identity providers (e.g., google, yahoo, facbook, etc). The outside users of these sites now authenticate at one of those sites, and those sites provide the local application with information about the browser user.

Since the mid-1990s commercial Internet-based Service Providers have allowed people visiting their sites to "sign up" and obtain an account. Almost always these accounts have the user supplying a userid (which must be unique within the site) and a password (which sometimes must meet certain strength requirements). Sometimes the userid is actually the user's email address at some other site. Oftentimes, the site asked the user to provide other information as part of their "profile". Several of these items would usually be classified as PII (eg name).

In the mid-90's, investors thought that attracting large numbers of people to create local accounts was a sign that a site was prospering and succeeding. Over time, though, users began to push back on creating so many accounts. In addition, protocols began to emerge to support Web Single SignOn across a multitude of commercial Service Provider sites. This reduced the number of places where a user would have to maintain an "account". In recent years, a handful of large sites have come to hold the vast majority of user accounts (ie google, yahoo, twitter, msoft?, other). Taken together, these sites are often described as providers of "social identities". These sites are currently using a variety of proprietary protocols (and variations of OpenID) to provide SSO functionality. The consensus is that it is extremely unlikely the social identity providers would agree to issue SAML Assertions to SAML SPs.

One significant difference between enterprise identities and social identities is that enterprise identities have been through a business process to ensure that the Credential is given to the physical person with whom the account is associated. Social Identities do not currently have an equivalent process. Users go to these sites and create accounts, and then self-assert profile information (eg a name). Consequently, social identities are not currently viewed as having a sufficiently high Level of Assurance to allow them to be used for business transactions where account information is linked to a person's name or real world identity. They can be used for business transactions when all the information needed to complete the transaction is entered by the person (eg a credit card number and validating information). They cannot be used for business transactions when a browser user is claiming to be a specific individual (eg accessing a government services site).

In the future, there may be processes provided by third parties to verify the real world identity of a person using a social identity.

See TN - Protocols Used by Social Identities for detailed information about the protocols used by Social Identity sites.

Issues for Management

Campuses are seeing a growing number of situations where application owners want their site to be used by authenticated users who cannot be authenticated by the home campus. There is a broad range of use cases and requirements in these situations; GenericUseCases provides descriptions of many of these. Some of the key issues that must be considered when an application owner (and, the campus IT organization, if they are involved) include:

1. Will most of the usage be by campus members, or by people from outside the campus ?
2. How important is it that information about a person's "real identity" be available? If a person edits content at the site, how confident do others want to be about the real identity of the editor ?
3. Does the site want to grant more permissions to an identity that is thought to have a higher LoA ?
4. Does the campus want to "remember" social identities that are being used to access sites on the campus ? Or is the campus comfortable with delegating this decision to individual sites?

Often, at the outset, the answers to these questions are not clear or obvious. And, often, the answers will evolve over time. And, as campuses move toward outsourcing some of their core services (eg email to Google or Microsoft), depending on how the campus chose to handle identity and authentication, it may be easier to rely on SSO provided by the outsourced provider than on SSO provided by the campus infrastructure. Lastly, an application site based in an academic department may have different answers to these questions than the central IT organization at that same campus.

(These notes assume that a site needs to support access via local, Federated, and social identities. The target audience for these notes is management in the campus central IT department.)

Comparison of Enterprise and Social Identities

Campuses issue digital identities to members of their core communities (faculty, students, staff); recently, they are also issuing credentials to communities that have "some" relationship to the campus, but a weaker relationship than members of the core community. The business processes supporting credential issuance to members of the core community require the person to prove their legal identity, and then associate that identity with a digital identity (a process known as "identity proofing"). As a result, the campus trusts that when some authenticates to a system using Jane Doe's credentials that it MUST have been Jane that performed this action. If misuse occurs during a session that was initiated with Jane's credentials, then Jane is held responsible. However, because many of the people now receiving credentials are remote from the campus, these assumptions do not apply to them.

In recent years, campuses have begun to leverage Federated Identity, and the InCommon Federation, in order to allow people from other campuses to access local applications and collaboration sites. The Federated Identity approach allows a campus to leverage the identity proofing processes used at other campuses. Event though the user is authenticating at a different site, the local campus trusts that the other campus has done an effective job of identity proofing before issuing digital credentials to people.

More recently, more and more application owners want their sites used by people from outside of the Higher Education environment. This group of applications ranges from the expected submitting comments on a blog to participating in a wiki supporting a research project to real business applications (e.g., a student giving their parents access to the student bill). Many of these sites are looking to have these outside users authenticate at the big internet identity providers. This effectively offloads most of the burden of supporting these accounts to the identity provider. The cost of providing this support is often non-trivial, so this represents a big win for the application owner.

However, currently no organization provides any sort of identity proofing process for these accounts. Though many people use their name or nickname as their account name at these providers, there is no requirement that people follow this informal convention. Nor is there a process that prevents someone from using another person's name or nickname. Consequently, there is no way of automatically obtaining trust about the identity of a holder of a social account.

Some applications attempt to raise their confidence level about the true identity of the social account holder by using out-of-band approaches. Talking to the person and asking then to share their social identity is a popular approach. This has the obvious drawback that the application owner must know the outsider in order to ask this question. Other sites include a signup or apply page; the site owner reviews an application, perhaps contacts the applicant, and decides whether to grant admittance and which privileges to grant. In this case, the application owner has no recourse beyond revoking privileges if the applicant begins to behave badly. A third approach is to allow current members to use the application to send invitations to social identities; the social account holder clicks a url in the email message in order to join the site and gain privileges. This approach is useful when the current member is allowed to share access to specific information with other parties (e.g., a student giving their parents online access to a student bill). The campus does not care who the student shares the bill with; that is the student's responsibility. The campus does not need to know the real identity of the social account holder.

In our call of 28 March Dedre said she might draft some "language differentiating social loa 1 from campus loa 1" ... if memory serves, Heather was also interested in contributing to this. While the text above this note describes how the processes for proofing identity differs between social and campus identities, the way I recall this call suggested that Dedre had in mind something that was grounded in use cases – like the several enumerated in the prior paragraph, but perhaps more elaborately illustrated. If this draft is floating around somewhere, and its focus is what I'm imagining/recalling, might it be the bones of a section describing risks of using social identity providers? (~Steve Masover, 27 Aug 2011)

--- what are the risks of using social identities

Possible Deployment Models

Several possible deployment models are possible, depending on the level of use, the amount of risk that the application owner and site are willing to accept, and the need to maintain a longer term relationship with the social identity holder. This list is ordered from simplest to more complex.

Isn't it true that the ordering here -- "simplest to most complex" -- is from a campus central IT perspective? It is true that this is the declared audience of this page, but perhaps a more nuanced presentation would be helpful here. What's simplest for central IT (#1) is hands down most difficult for an application / site owner. (~Steve Masover, 27 Aug 2011)

1. The application owner deploys a Social to SAML gateway as part of their site. The site owner is responsible for all the work and configuration. There is no involvement by campus central IT.
2. The application owner uses a Social to SAML gateway operated by some other party, outside of the campus.
3. Campus central IT deploys a shared Social to SAML gateway. An application on the campus is allowed to leverage the gateway.
4. Campus central IT deploys a shared Social to SAML gateway, and elects to "remember" all of the social identities that cross the gateway.

Implementation Process

(for each option, guesstimate the level of effort and required skills)

-- existing open source implementations

-- existing commercial implementations

-- issues for the "build it locally" approach

-- comparison of enterprise and social identities
  -- pro's, con's and risks of the two approaches
  -- quick description of possible deployment models
  -- where to get the pieces (buy, open source, build)
  -- how much effort
  -- what are the risks of doing this ?

  -- other topics ?

Suggestions for Implementers

Perspectives on Handling Both Social and SAML Identities
Developers
Gateway Developers
Deployers

Defining the Problem, and Identifying Approaches

Increasingly, groups and people operating web sites on campuses want their site to be used by members of their campus community, by people from other campuses, and by people from outside the Higher Education/Research environment. Central IT Departments have provided authentication mechanisms that address the first two groups. The last group is often addressed by adding user management (self-service?) and authentication to the application; however, more recently, web sites have moved toward relying on social identity (provided by the big Internet providers). Web site developers and deployers are now looking for ways that their applications can support both enterprise and social identities.

1. Perspectives on Handling Both Social and SAML Identities

What are each of the parties hoping that an approach, a framework, will provide for them ?

- Relying party wants to make some federated resource accessible to people who have a Social Identity Provider (Twitter, Yahoo, Google, Windows Live)
  - On one hand they want to minimize changes to their Service Provider implementation and their application code
  - On the other, they want to know which ~~and what type of~~ IdP is handling any particular access instance

    - They want the information about an authentication event presented to them in a standard format, independent of the protocol that was used. If they want identity attributes as well as an authentication assertion, they want those attributes to have consistent names and consistent value syntax. The same attribute name should not represent two different attributes nor should the value syntax vary for a given named attribute. Different implementations providing social identity support should use standardized syntax and semantics for all the provided attributes.
        - This applies whether a Social-to-SAML gateway is involved or not.
    - They do not want to be bothered with designing and implementing a complex Discovery process that can accommodate multiple protocols. They would rather hand this problem off to a gateway. However, this also offloads part of the problem onto the user experience.
    - They want the application to contain a mechanism that allows a current user to "invite" a new participant to join the site, independent of the type of identity that individual might eventually use to authenticate to the site.
    - They want the application to contain a mechanism that allows a new interested individual to join the site, or apply to join the site.

- The user expects to make a search-and-one-click selection of their IdP of choice
    - They would like to see a given social identity provider identified the same way regardless of their path to the SP
    - They might expect that they would be recognized as the same individual regardless of their choice of IdP, but in general this is not possible without some user mediated account linking on a per-SP basis

1.2 Models for Integration

- A native SP implementation supporting both SAML and Social protocols that deployers could add to their site.
    - unfortunately, such an implementation does not exist (really? simplesamlphp ?)
    - Currently, there is no single "standard" protocol used by the various social identity providers. They seem to use proprietary protocols or proprietary variants of OpenID. Change happens quickly, and is outside the normal standards processes. This churn means that groups that maintain native SP implementations will often wait for some level of consensus before providing support for a newer protocol.
    - Going forward, there seems to be broad interest in both oAuth V2 and OpenID Connect. If broad deployment and consensus appear, native SP implementations may add support.
- A central gateway, operated by a Federation, that translates incoming protocol requests to a single standard protocol which is supported by all of the SPs in the Federation.
    - unfortunately, due to the frequent churn in the protocols used by the social sites (and the fact that some of these protocols are proprietary), no Federation is willing to run such a gateway (and have to deal with the long term support of such a GW). (Actually, no -- FEIDE is doing this with Roland's GW)
- a gateway operated by a campus that translates incoming protocol requests to a single standard protocol which is supported by all of the SPs on the campus. It would be the campus' decision as to whether to limit the use of such a gateway to SPs located on the campus.
- a gateway installed by, maintained by, and operated by the SP site that supports translating incoming protocol requests to a single standard protocol which is supported by the application.
    - note, though, that the reason SP operators have been looking to use gateways is that they want to offload the responsibility for the Discovery process to the gateway.

1.3 Gateways as Necessary Evils--For a Time

- The majority of current identity federation deployments connect SAML IdPs with SAML Relying Parties (RPs, SPs)
- A growing number of R&E organizations wish to make some of their services accessible to users who prefer to use Social Identity Providers (Twitter, Yahoo, Google, Windows Live, Facebook). Some of these users will in fact not have an R&E organizational SAML identity.
- R&E application owners do not want to modify each SAML-protected SP to handle social identities
- As a result, many have been attracted to the idea of a Social-to-SAML gateway that permits a user to authenticate with their social IdP of choice and access a SAML relying party application or service
- The Social-to-SAML gateway is often designed to take the role of SAML IdP in interactions with RPs. This is driven by the application owner constraint that the solution not involve major modifications of existing RPs and SPs
- So a key function of the Gateway is to map social IdP asserted authentication events and identity attributes to a corresponding SAML assertion for consumption by the RP.

1.4 The Proposed Model

-- the SP will have to include support for Discovery

-- by configuring their Discovery mechanism, the SP decides which social providers to allow and trust. if the SP wants to exclude facebook, then the GW has to be able to enforce that policy choice....

-- the SP would also configure the address for the social-to-SAML gateway that they choose to use.

-- the browser user would select their identity Provider (eg a campus, google, yahoo, facebook, etc), and click SUBMIT.

-- the user would be redirected to the gateway; the SP would send an entityID value that identifies the requested social IDP)

-- the browser user would be redirected on to the social provider, authenticate, and be returned to the gateway.

-- the gateway would construct a response to the SPs AuthnRequest, using the guidelines described below.

-- the gateway would issue a POST back to the SP, using a standard SAML flow.

NOTE -- in this model, the user does not see the gateway ever present a GUI

1.5 Basic Gateway Usage Model

1. The browser user would select their identity Provider (eg a campus, google, yahoo, facebook, etc), and click SUBMIT.
2. The user would be redirected to the gateway; the SP would send an entityID value that identifies the requested social IDP)
3. The browser user would be redirected on to the social provider, authenticate, and be returned to the gateway.

4. The gateway would construct a response to the SPs AuthnRequest, using the guidelines described below.
5. The gateway would issue a POST back to the SP, using a standard SAML flow.

NOTE -- in this model, the user does not see the gateway ever present a GUI

Case Studies

- Penn State OpenId Implementation - mod_auth_openid wrapped around a Shibboleth IDP
- Implementation Descriptions

2. Developers

This section is intended to be a cookbook-like document for people developing applications and who want to allow authentication from social identity providers (eg the Bamboo project)

Suggestions

1. Isolate your application from the authentication protocols.
2. Choose an authentication package implementation that supports standard attribute conventions. See TN - Conventions on Attributes
3. The application would need to know both 1) the identity of the social identity provider, and 2) the identity of the gateway which is forwarding the authentication event in order to determine whether or not to trust the presented Assertion.
4. The SP should include support for the Discovery Process. By configuring their Discovery mechanism, the SP decides which social providers to allow and trust. if the SP wants to exclude facebook, then the GW has to be able to enforce that policy choice....
5. the SP would also configure the address for the social-to-SAML gateway that they choose to use.
6.
7. An application will need to map incoming identities to same internal "person object"

3. Gateway Developers

How to Build an Eventually Dispensable Gateway Service

- The goal is eventually to outgrow the need for gateways by providing native multi-protocol SPs that support both SAML and selected Social IdPs.
- This goal will be easier to achieve if the gateway functions are essentially invisible to the end user. That is, the step in which a user selects an IdP should behave the same way whether or not a Social-to-SAML gateway is involved.
- Unique EntityIDs should be defined for each SocialIdP-Gateway pair. This allows SPs to decide on trust points based on information in SAML2 metadata files if the relevant identity federation supports this.
- Mapping social IdP assertions to SAML assertions is a core gateway function

Gateways Must Honor SP Policies on Acceptability of Various Social Providers

- If an SP does not accept a certain Social IdP (e.g., they opt not to accept Facebook-based authentications), the gateway must honor that policy by not offering the prohibited IdP in the IdP selection list available to the end user.

could define policies at the SP saying who is allowed to assert which kinds of identifier; once you make it thru that layer of filtering than the app can trust it...

eg if you trust GWs X, Y, and Z, and IDPs 1, 2, and 3

4. Deployers

1. Because the SP configures in which GW it wants to use, it doesn't have to worry about people coming in from multiple GWs and using the same social provider.
2. could define policies at the SP saying who is allowed to assert which kinds of identifier; once you make it thru that layer of filtering than the app can trust it...

eg if you trust GWs X, Y, and Z, and IDPs 1, 2, and 3

Policy Issues to Consider

1. After reviewing your goals for your site, and the types of data that it will contain, identify which social identity providers 9and associated protocols) you want to trust.

Gateways Must Honor SP Policies on Acceptability of Various Social Providers

- If an SP does not accept a certain Social IdP (e.g., they opt not to accept Facebook-based authentications), the gateway must honor that policy by not offering the prohibited IdP in the IdP selection list available to the end user.

could define policies at the SP saying who is allowed to assert which kinds of identifier; once you make it thru that layer of filtering than the app can trust it...

eg if you trust GWs X, Y, and Z, and IDPs 1, 2, and 3

Account Linking: A process by which a person links multiple modes of authentication (i.e., authentication through multiple Identity Providers) to a single logical identity in an application or system (e.g., a single User ID that refers to a single person). "Account Linkage" is defined in the SAML 2.0 glossary as: "A method of relating accounts at two different providers that represent the same principal so that the providers can communicate about the principal. Account linkage can be established through the sharing of attributes or through identity federation."

Assertions: An assertion is a unit of information related to security. Assertions are made up of statements that can be used to make decisions about authorization (access to resources). The definition of "Assertion" in the SAML 2.0 glossary is: "A piece of data produced by a SAML authority regarding either an act of authentication performed on a subject, attribute information about the subject, or authorization data applying to the subject with respect to a specified resource."

Authentication (AuthN): The process of proving one's identity to an application or system. Often this involves presentation of "credentials" – e.g., a User ID and a password. More broadly, a person is authenticated based on "factors of identification" that may include something the person knows, something she has, or something she is. Examples of each of these: a person might *know* a password; might *have* an identity card; and might *be* identifiable based on biometrics such as a fingerprint or retinal pattern. When a user is authenticated, her identity is considered to be known to some level of certainty or assurance. The definition of "Authentication" in the SAML 2.0 glossary is: "To confirm a system entity's asserted principal identity with a specified, or understood, level of confidence."

Authorization (AuthZ): The grant or specification of access rights to resources within an application or system. For example, a doctor may be *authorized* to view her patients' medical records. The definition of "Authorization" in the SAML 2.0 glossary is: "The process of determining, by evaluating applicable access control information, whether a subject is allowed to have the specified types of access to a particular resource. Usually, authorization is in the context of authentication. Once a subject is authenticated, it may be authorized to perform different types of access."

Collaborative Organization: (cf. *Virtual Organization*)

Credentials: Evidence presented in the course of an Authentication process intended to prove identity. A familiar form of credentials is a User ID and the Password associated with that User ID. "Credentials" is defined in the SAML 2.0 glossary as: "Data that is transferred to establish a claimed principal identity."

Digital Identity:

Discovery:

Enterprise Identity:

EntityID:

Exposure:

Federated Identity: A mode of establishing identity that is agreed-upon by multiple identity and/or service providers. "Federated Identity" is defined in the SAML 2.0 glossary as: "A principal's identity is said to be federated between a set of Providers when there is an agreement between the providers on a set of identifiers and/or attributes to use to refer to the Principal"

Federation of Trust: (cf. *Trust Federation*)

Gateway: (cf. *Social-to-SAML gateway*)

InCommon Federation: The InCommon Federation defines itself as "a formal federation of organizations focused on creating a common framework for collaborative trust in support of research and education."

Identity Proofing: Assuring that digital identities and associated credentials belong to the real-world individuals they purport to represent. Cf., for example, *What Is Online Identity Proofing and How Does It Work* (eWeek, 2 Aug 2007)

Identity Provider (IdP): A type of application or system that is trusted to authenticate users. "Identity Provider" is defined in the SAML 2.0 glossary as: "A kind of service provider that creates, maintains, and manages identity information for principals and provides principal authentication to other service providers within a federation, such as with web browser profiles."

Level of Assurance (LoA):

LoA: (cf. *Level of Assurance*)

OAuth:

OpenID:

out-of-band:

PII: (see *Personally identifiable information*)

Personally identifiable information: Information that can be used to identify a particular person. Examples of personally identifiable information (sometimes referred to as "PII") include: full name, national identification number (e.g., passport number or social security number), driver's license number, credit card number(s), birth date, birthplace, etc.

Relying Party: An application or system that relies on another application or system to fulfill some of its functionality (e.g., to authenticate a user). "Relying Party" is defined in the SAML 2.0 glossary as: "A system entity that decides to take an action based on information from another system entity. For example, a SAML relying party depends on receiving assertions from an asserting party (a SAML authority) about a subject."

Risk:

SAML / **SAML2**: SAML is an OASIS standard that defines how security-related statements (assertions) are expressed; cf. SAML 2.0 on the Oasis Standards page; and the OASIS Security Services (SAML) Technical Committee page. As defined in the SAML 2.0 glossary , "Security Assertion Markup Language (SAML)" is: "The set of specifications describing security assertions that are encoded in XML, profiles for attaching the assertions to various protocols and frameworks, the request/response protocol used to obtain the assertions, and bindings of this protocol to various transfer protocols (for example, SOAP and HTTP)."

SAML2 metadata files:

Service Provider (SP): An application or system that delivers functionality (services) to people or other applications / systems. As defined in the SAML 2.0 glossary , a "Service Provider" is:"A role donned by a system entity where the system entity provides services to principals or other system entities."

Single SignOn (SSO):

Social Identity:

Social-to-SAML Gateway:

Trust Federation: An organization (such as the InCommon Federation in which members trust each other to honestly and reliably authenticate users and present information about them ("assertions") to others in the federation.

Virtual Organization:

Web Single SignOn (SSO):

A set of use cases has been submitted that describe campus central IT adding "social identity people" to the central person registry (and perhaps associating multiple sets of credentials with an individual); a separate set of use cases has been submitted which include no role for central IT or for "remembering" anything about the person using social credentials. Both models seem to have significant numbers of people interested in them. Consequently, both models are likely to be deployed, with campuses choosing a model appropriate to the problem they are trying to solve.

One of the institution level models involves imposing some control over which social identities are allowed access to services (perhaps by requiring an invitation).

Perspectives on Handling Both Social and SAML Identities

- Relying party wants to make some federated resource accessible to people who have a Social Identity Provider (Twitter, Yahoo, Google, Windows Live)
    - On one hand they want to minimize changes to their Service Provider implementation and their application code
    - On the other, they want to know which ~~and what type of~~ IdP is handling any particular access instance
    - If they want identity attributes as well as an authentication assertion, they want those attributes to have consistent names and consistent value syntax. The same attribute name should not represent two different attributes nor should the value syntax vary for a given named attribute.
        - This applies whether a Social-to-SAML gateway is involved or not.
- The user expects to make a search-and-one-click selection of their IdP of choice
    - They would like to see a given social identity provider identified the same way regardless of their path to the SP
    - They might expect that they would be recognized as the same individual regardless of their choice of IdP, but in general this is not possible without some user mediated account linking on a per-SP basis

The issue that has too many names: Invitation, Volunteering, Conscription and other ways of adding members to a CO

All collaborating organizations need collaborators. How do individuals get a CO membership? Process models proliferate to cover all the ways this needs to be done. We identify the basic modes and discuss their applicability, their strengths and some of their unresolved issues.

Invitation

Volunteering

Conscription

Google, Facebook, Twitter, Windows Live and More: How Social Identities Work

Links of Interest

Google

Federated Login for Google Account Users
OpenID Federated Login Service for Google Apps
Authentication and Authorization for Google APIs

Facebook

Authentication
Facebook for websites

Twitter

Overview of "Sign in with Twitter"

Windows Live

Introduction to Windows Live ID
Windows Live Developer Center

Other articles or presentations of interest

Intro to IdM for VO (PPT)

1) There is a need for a formal process to assess the severity of the risk and exposure that results when an SP is accessed using a stolen credential.

The existence of this process, the model it is based on, and the process for using it to evaluate a specific SP needs to be communicated to IT staff, business owners, and the security group.

-- education needed for everyone (IT staff included) that just because two systems use passwords doesn't mean they're comnparable on LoA

-- important role of auditors was noted

2) Do campuses currently have any relevant policies or frameworks ?

-- campuses have policies related to abusive behavior; some campuses have data-driven policies (ie data owner responsible for classifying data)

3) Are there other approaches in wide use to assessing risk at an SP ?

-- OMB 04-04 describes a framework for assessing the impact at an SP of an "Authentication Error" (ie someone using a stolen identity).

http://www.whitehouse.gov/sites/default/files/omb/memoranda/fy04/m04-04.pdf

which is the foundation for the US Federal government's work on risk and Levels of Assurance.

Section 2 contains useful suggestions on how to assess risk. It identifies several categories of harm or impact:

Inconvenience, distress or damage to standing or reputation
Financial loss or agency liability
Harm to agency programs or public interests
Unauthorized release of sensitive information
Personal Safety
Civil or criminal violations

For each of these categories, it provides a definition and examples of low, moderate, and high impacts.

It also provides guidance on mapping low, medium, and high risk in these categories to an appropriate required LoA.

Table 1 – Maximum Potential Impacts for Each Assurance Level (from OMB 04-04)

Assurance Level Impact Profiles				
Potential Impact Categories for Authentication Errors	1	2	3	4
Inconvenience, distress or damage to standing or reputation	Low	Mod	Mod	High
Financial loss or agency liability	Low	Mod	Mod	High
Harm to agency programs or public interests	N/A	Low	Mod	High
Unauthorized release of sensitive information	N/A	Low	Mod	High
Personal Safety	N/A	N/A	Low	Mod High

Civil or criminal violations	N/A	Low	Mod	High

4) How is the NIST work applied to this situation ?

Campuses and Service Providers could use the NIST framework to evaluate, for a specific SP and for each category, whether an "Authentication Error" would create a Low, Moderate, or High Impact situation. They would then use Table 1 above to determine the minimum LoA that is required of an Authentication Event in order to access the SP.

5) Is the NIST work relevant to this problem ?

The Consensus is YES, but that its not sufficient.

social identities are at LoA 1 (as decreed by the US government).

Most campus identities are also (currently) at LoA 1 (also referred to as InCommon Bronze; sometimes called "whatever the campus is doing today").

However, there was consensus that a campus-asserted identity was "stronger" than a social identity. Currently, no attempt is made to link a social identity to a real world identity. Campuses do make this effort, and promulgate policies prohibiting the sharing of identities and passwords. Campuses may not be operating at NIST Level 2 (InCommon Silver), but do have have business practice and policy in place to create a "reasonable" assurance about who is using a specific credential.

Implementation Issues

6) LoA is orthogonal to differentiating social vs enterprise identities. SPs won't implement complex algorithms, just a couple of differentiators at this point. Initially, avoid the temptation to provide an SP with too much information.

7) Consensus -- GW should assert both authN source and forward any LoA value that it receives; the application can use application context + whatever algorithm it wants to determine how it wants to treat an incoming Assertion (eg any SP can decide that social + LOA 1 --> treat as LOA 0)

8) Does the GW compute an LoA value ? Should the GW differentiate social, campus bronze, silver in some sort of LoA assertion that it computes ?

Consensus -- NO. Social identities are LoA 1; GW forwards any eduPersonAssurance attributes or Bronze/Silver assertion that it receives. But, it does NOT attempt to compute or derive such a value.

.. as long as the operational characteristics of the GW don't impact LoA that is being presented.

9) some applications might want to know some of the properties of the authentication event, separate from LoA (eg was google 2-factor used)

Let any further categorization emerge from experience and practice...

10) There's also a need to ensure that enough logging is being done, so that forensics are possible, and a campus/site manager could take an SP out of service quickly, do the required investigation and recovery/correction, and return the site to service.

This section contains a set of recommendations on "standard" definitions for interfaces and data encoding across the boundatries where the various components must interact.

1. Attributes

Native SAML implementations are already using standard interfaces between web servers and application to provide the information that has been asserted by an IDP. In addition, these implementations are also leveraging a standard set of attribute syntax and semantics. These also provide a framework that can be easily extended to accommodate new attributes. The more consistent the Social-to-SAML mappings are across gateways, the easier it will be to transition away from gateways to native multi-protocol SPs.

1.1 Principles

1. Everything must be expressed in a manner that is independent of whether a native or GW implementation is being used.
2. Different social identity providers should be represented by different uri values (this is the same convention used with every other saml idp provider). This is essential for implementing Discovery in a seamless fashion. Its not necessary that every GW in the world use the same identifier for each social identity provider. Clearly, though, consistency would a good thing.
3. A given person's identifier from a given social IdP should be the same, regardless of which gateway it passed through
4. Since different Social IdPs may use different attributes to carry the user identifier, the gateway should use different attribute names to carry a given Social IdP's user identifier.
5. The SP should be able to determine which Gateway (if any) processed the request. (eg an SP might trust one GW but not another)
6. The SP should be able to determine which social IDP authenticated the user and issued the initial assertions.
7. The various social IDPs assert a variety of PII information; different providers assert different information. Any asserted PII should be mapped to the appropriate ldap/SAML attribute. Many of the collaboration sites which would use this support want the user to present several PII-attributes (eg name, email, identifier for this principal from the social provider). However, not all social authentication providers will share this information (especially email). As a result, the application may have to present a "user profile" form the first time a user connects, and ask the user to self-assert various values.
8. Consistency is desirable in several areas
    a. Organization Identifiers--or Why EntityID was Guaranteed to be Misused

1.2 Gateways Mapping Social-IDP Provided Information to SAML Assertions

(similar to a profile, set down some specific rules)

1. The entityID value asserted by a Gateway as the Issuer MUST be a concatenation of (a gateway identifier) + (the social identity provider), since different GWs might use different values to represent a specific social identity provider.
2. The resulting entityID will NOT be parseable, but will be mappable (mapping file)
3. The value in the Subject/ NameID element MUST be the user identifier asserted by the social IDP. (eg a google account from one GW MUST be the same as a google account from another GW).
    a. NOTE For instance, a social-SAML gateway, or a Virtual Organization, could assert a google user identity. If they both both asserted a user identity of X@ google, presumably it is the same person being asserted by two different entities. NOTE -- accounts are different from email addresses, even though with some social providers there might have the same value).
4. The identity of the social IDP MUST be passed in the XXX attribute, and encoded as follows:
    a. Google -
5. The identity of the processing Gateway (if any) MUST be encoded as the YYY attribute
6. The identity of the browser user MUST be encoded as an attribute, and using different attribute names to represent identities coming from different social providers; using a different uri gives you the opportunity to tell GW implementers what to do
7.

We need to do an appropriate study of what all the identifiers in the social space are -- we might end up with different attributes from different providers, or different syntax in some cases.

Future Proofing SAML Assertions Containing Social IdP-Provided Information

- There will be any number of gateways in use in the near term

2. Discovery

(some text about the rules between the SP and the GW, when it comes to Discovery related issues)

Use Cases

A common use case

An institution (virtual or traditional) creates a service that is designed to maximize collaboration and/or participation. The individuals participating in this service do not necessarily stay associated with any one institution over the lifetime of their participation. For example, a graduate student participating in a wiki space using an account with a strong LoA graduates and becomes a postdoc at another institution, still within the same field and expecting to remain active on this service. To bridge the use of the identity from the old institution with the identity from the new institution, the individual starts to use their Google/Facebook/other social identity account to continue to access material. Associating the old identity with the social identity and later with the new identity needs to be something the individual controls.

A set of generic use cases

Interesting Questions/Topics to Comment on.

The iPlant Collaboration wants to limit what people coming in from social identity providers can do as compared to people coming in from federated identity providers. However, iPlant, like many other VO, does not want to have to do the manual reconciliation of identity when people move from institution A to institution B. One idea, then, is to have that social identity be the bridge to link institution A's account to institutions B's account. However, having someone with limited permissions because they are coming in from a social identity provider have the power to link higher "value" federated identities... that's just not ok.

Then there's the idea that a person can link the account after they get to their new institution. But for arguments sake, one can say that they are coming from an institution that actually has a reasonable account lifecycle policy (don't laugh - this will be the norm some day) and their old account is no longer accessible in any way for linking. No authentication at both institutions simultaneously is possible. All this feeds back to manual reconciliation on the part of the VO or whoever is handling the identity linking in the service, and they don't have the resources to manage this.

Some faculty members have dual, maybe even triple appointments across multiple institutions. At Stanford University, a doctor at one of the two hospitals will have an account at one (or both) hospitals, and will automatically have an account at the School of Medicine as faculty. And, since all of this falls under the Stanford umbrella, they may well have a Stanford id as well. When they join a collaboration, the doctor/faculty member will not think about which id they signed up under, they will want all id's to have equal value and access to the same information.

UseCase 2: Institutional ID linked to Social Identity – Should Institution Allow sign on via Social Id?

- An institution creates an id locally and adds it to it's identity ecosystem.
- Said institution permits account linking (inbound, from Social Id to Institutional one)
- Should institution allow end user to sign into Social ID for regular account access?
    - Why or why not?

UseCase 3: Institutional ID (student) linked to Social Identity (parent)

- An institution creates an id locally (for a student) and adds it to it's identity ecosystem.
- Said institution permits account linking (inbound, from Social Id to Institutional one)
- Can this same scenario (UseCase2) be used to provide access to a limited number of resources using the Social Identity of the parent?
    - Why or why not?
    - The resources MAY require a higher LoA to access (e.g. student account, class schedule, etc.)

Style of linkage	Pro	Con	Comments

Inbound into institution	Ease of use for end user	password strength requirements at mercy of Social Id	This may not be the only risk