Multiple Identity Providers and Level of Assurance

Author: Nicole Harris (JISC)
Contributors: Andrew Cormack (JANET(UK)), Licia Florio (TERENA), David Chadwick (Univ Kent), Alex Reid (AARNet).

Abstract
This document is a summary of a debate on the TF-EMC2 mailing list in February 2009. The discussion regarded the introduction of non-institutional Identity Providers (IdPs) into HE Federations and the impact that this would have on end-user choice, end-user processes and securing levels of assurance. This document it is intended to inform future discussion and developments in these areas within TF-EMC2, REFEDS and the wider community.

Table of Contents

Overview

This document is a summary of a debate on the TF-EMC2 mailing list regarding the introduction of non-institutional Identity Providers into Federations and the impact that this would have on end-user choice, end-user processes and securing levels of assurance. It is intended to inform future discussion and developments in these areas within TF-EMC2, REFEDS and the wider community.

Four core questions were raised in the debate:

1. Should non-institutional authentication systems (Facebook and similar) be allowed as Identity Providers (IdPs) within Federations?
2. Should users have free choice as to the Identity Provider they use?
3. What is the impact on Levels of Assurance of allowing multiple Identity Provider routes for users?
4. How do we aggregate attributes from multiple sources for authorisation purposes?

Answers to these questions will be provided in the following paragraphs.

1. Should non-institutional authentication systems be allowed as Identity Providers (IdPs) within Federations?

The discussion on this point was triggered by the proposal to allow SPs in the FEIDE[1] federation to add Facebook as IdPs. The rationale for this was the need to lower the barrier to access to FEIDE. It is somehow perceived that access to the identity federations (both for SPs and end-users) is still rather difficult.

IN FAVOUR	AGAINST
F1. Federations should promote 'openness' and not be closed clubs.	**A1**. Should not be allowed where the authentication system is specifically linked to a system that does not allow openness itself (i.e. does not accept SAML authentication).
F2. Institutions are looking to extend their offerings to include access routes via OpenID and other similar processes in conjunction with institutional infrastructure. The process is seen as a natural expansion of federations, which moved from providing access to educational services, to also include commercial SPs and is now considering commercial IdPs.	**A2**. Could be seen as promoting alternative routes and downgrading the importance of the institutional IdP. Particularly concerning users' privacy, federations are much more aware and preserve users' privacy better than commercial companies like Facebook.
F3. Multiple identity provision reflects the real world and can be compared to the affiliations we carry around in wallets. Furthermore the user-centric approach will become more and more popular in the near future, so federations should get ready for	**A3**. Federations should not promote the sharing of user data with private companies.

this.	
F4. Allowing multiple IdP routes promotes the future model of attribute aggregation.	**A4**. Trust in the IdP is the key factor within federations and commercial IdPs do not offer the right level of trust or it is unknown or variable.
F5. Providing all they do is authenticate their users, and do not provide any authz attributes (since these are self-asserted), then this is a useful service to perform. The SP can be assured (to some level) that it is the same user every time, even if it does not know anything else about the user.	**A5**. Very few services will be happy in not knowing anything about the user other than it is the same one as last time, and depending on the external IdP, they may not guarantee non-reusability of user Ids.

2. Should users have free choice as to the Identity Provider they use?

The discussion around this point related to the fact that there can be a separation between who provides the actual credential used (an OpenID or an institutional identifier etc.), who validates that this belongs to a real world person, who adds attributes that recognise this person's affiliation(s) and who authorises access to a service based on all of this information. There seems to be consensus on the fact that Identity vetting is a non-trivial task, so there might be benefits having a third party handling it.
In this view, if Google after offering e-mail services to students would be willing to take a next step and offer identity proofing services that meet the university's operational and privacy requirements universities might be able to delegate the identity proofing to Google for instance.

The importance of how all of this information is revoked was also highlighted.

In current Federation models, the institution tends to provide identity assurance, provisioning of identity credentials and attribute management. Disaggregation of these roles would introduce interesting new challenges for institutions.

For attribute management to work effectively in these scenarios, an attribute schema that goes beyond the limits of the "edu space" would need to be developed.

There is also a complication about the interaction between publicly funded NRENs and commercial services such as Facebook.

Is identity provisioning / vetting a core task for institutions? Who do I as a user trust? Who does a Service Provider trust?

3. What is the impact on Levels of Assurance of allowing multiple Identity Provider routes for users?

The need to allow users to have multiple identity management routes, and the need for attribute aggregation across these routes is a well recognised problem within the federated

access space. Such attribute aggregation will necessarily have an impact on levels of assurance in the user credentials, particularly where different routes offer different levels of assurance – such as an institutional Identity Provider versus a commercial, user-managed credential.

Levels of Assurance (LoA) can be added to an identity at several stages, but are most commonly recognised as Registration LoA (robustness of identity vetting at registration) and Authentication LoA (strength of credential). Combined, these could be described as creating the Credential LoA. Note that if an IdP only provides an authentication function and does not provide any user attributes (other than the authn subject identifier which is typically meaningless), then the Registration LoA is of no concern to the IdP since the Registration LoA is the assurance that is given to the registered attributes (and in this case there are none). In this scenario the Credential LoA should equate to the Authentication LoA. If on the other hand this IdP did allow user self-asserted attributes to be created at registration time, then the Registration LoA would be zero, and so should the Credential LoA applied to the authz attributes.

If attribute aggregation occurs across multiple identities, it is sensible to review whether the Credential LoA can be increased by improving the Registration LoA at a later date. A user could create an online identity with zero Registration assurance (for example by inputting his own username and password on the New User Account page), and this could have a reasonably good Authentication LoA (for example by using a complex very long password transferred over SSL). If the user provided self-asserted attributes when creating his new account, then the Registration (and Credential) LoAs would be zero. The Registration LoA could then be increased from zero up to the Authentication LoA at a later date by the IdP undertaking thorough real-world identity verification of the user's self asserted attributes, and then applying this Registration LoA to the newly verified online identity. This would be a relevant area to examine for institutions that may consider brokering authorisation for resources around a student-managed OpenID[2] (as presented by David Chadwick at the 2008 Spring Internet2 meeting).

The chaining of LoAs from different providers causes problems as it is difficult to establish at what point which LoA applies, and indeed when it no longer applies. LoA requires both technical trust and procedural trust and when multiple parties are involved, this becomes complicated. This also creates a model where there can be multiple Registration Authorities for each set of credentials. The value chain for security is no stronger than the weakest link. If this implies that we need to know (and to a certain degree understand) the whole value chain, it is easy to remain committed to single Registration credentials since they by default (since they are well-known) offer better security. The LoA aggregation problem has been addressed by the Shintau project[3] at the University of Kent and a viable model (and implementation) has been produced, which will be demonstrated at the 2009 Spring Internet2 meeting, and described in the May 2009 edition of IEEE Computer Magazine.

A larger question revolves around who should mandate the level of security required for access. Identity Providers, Service Providers and Federations are all currently involved in this process, but there is no definitive answer as to where levels are mandated. (Ultimately is must be up to the SP to decide since it is its resources that are being used, though ideally there must be some negotiation, since the SP needs to take into account the ability of an IdP to provide any given level of security.)

The requirements of Service Providers as well as Identity Providers need to be taken into consideration. It can be argued that most Service Providers in the federated model are more interested in Registration LoA, since it is the registered attributes which are being used to provide authorisation: indeed in general where a Service Provider is making an access control decision based on an attribute value then their concern should be the Level of Assurance with which the attribute is asserted, whether this attribute is a personal identity or membership of a class such as 'student'. Registration LoA is also the most expensive process in the chain, so the financial impact on the strength of Registration LoA is important to both Identity Providers and Service Providers. The trusted third-party Identity Provider is crucial in the Service Provider acceptance of the federated model.

4. How do we aggregate attributes from multiple sources for authorisation purposes?

The Attribute Aggregation Model

The Attribute aggregation takes place at the Service Provider so the authorisation provider becomes the attribute aggregator. An example of this has been developed in the UK Shintau project using OASIS SAML attribute pull requests and the Liberty Alliance Discovery Protocol[4].

The general model the Shintau project employs is that at each IdP there is a registration phase, with a consequential Registration LoA, and an Authentication/logon phase with a consequential Authentication LoA. Different authentication mechanisms (PKI, Kerberos, OTP etc.) will have different Authn LoAs. Different IdPs may have different Registration LoAs for their attributes and different Authn LoAs for their authentication protocols. The current Session LoA is the LoA applied to the attributes aggregated in a particular session. The Session LoA is computed as the lower of the Authn LoA and Regn LoA at the IdP used by the user to login in for this session. Or to put it another way, the Authn LoA can never exceed the Regn LoA, otherwise you are asserting that someone's attributes are known to a higher level that you first registered them, which of course is impossible. Conversely, if one of the IdPs whose attributes are to be aggregated has a Registration LoA that is lower than the current Session LoA, then its attributes cannot be aggregated in the current session, otherwise they would be being asserted at a higher level than they were first registered (which again should be impossible). All the examples below fit into this general pattern.

Use Cases

Case:	**UC1: American Medical Schools (AAMC)**
Scenario:	The American Medicals Schools (AAMC) administers a test for admission into accredited US medical schools. Accounts are primarily given to users via e-mail verification to allow for the application process, but full identity proofing is then undertaken (fingerprinting and photo) when the students come to take the test. Campuses could benefit from capturing the value of the AAMC identity-proofing process.
LoA Details:	The initial Registration LoA is low (email verification only), which means that the Session LoA will remain low no matter how good the authentication mechanism is. After the students have taken the test, the Registration LoA is now high (due to fingerprinting etc.), so the Session LoA can rise to the lower of the Authentication LoA and Registration LoA.

Case:	**UC2: PGP web of trust**
Scenario:	Suppose I create a new, self-signed PGP key, with an e-mail address that you don't know. At that point you have zero confidence in the ownership of that key. Then we meet at the TERENA Conference and I say "I've got a new PGP key". At this point we can do whatever checks you feel you want in order to prove that I am the person I claim to be and that I hold the private key. After this you (should) have whatever level of assurance you want in the identity of the owner of the key and, because I have used an zillion-bit private key and you are confident that I'll take care of it, you also have sufficient level of assurance that if you later see a message signed with that key then it comes from me. If I had only used an eight-bit key then you might take the view that there was no point expending any effort validating my identity because someone else could easily reproduce the key pair.
LoA Details:	The Registration LoA is initially zero, because all the attributes were self asserted by some remote user you had never met. Once you meet face to face the Registration LoA increases dramatically. All the attributes in the PGP key now have a high level of assurance due to high Registration and Authentication LoAs (face to face and zillion bit key). If however the PGP key had only been 8 bit, then the Session LoA would remain close to zero, since anyone could reproduce the key pair.

Case:	**UC3: Students Using External Identities**
Scenario:	Suppose I create a user-centric AgorID identity for a username you do not know. And suppose the AgorID provider does no checks as to who I am in the real world: it just issues me with unique credentials that I can use to allow it to recognise me next time I come back. I then turn up as a

	student. You can do all the same checks as normal that I am the right person, have the right school exam results, have paid my fees, am entitled to be in the UK, etc. But in a fully federated world, you do not need to issue me with a new username and password: all you have to do is verify that the one I already own is of sufficient technical quality, and that AgorID's processes are sufficiently robust that it will never issue those same credentials to someone else, so that it doesn't undermine the checks you have done. And then (by your attribute linking process) assert attributes to service providers such as "student@here" that the original AgorID provider didn't know
LoA Details:	Initially the Registration LoA is zero due to self-assertion, and therefore so is the Session LoA. Once you turn up face-to-face and are verified then you can continue to use the AgorID only now with a higher Session LoA. (This is in fact the scenario presented by David Chadwick in the 2008 Spring Internet2 presentation.)

Case:	**UC4: Virtual Worlds**
Scenario:	CTO of a company providing virtual world service to teenagers. They are not at all interested in the real identity of their end users i.e. who they are in real life. There is a self-service registration with no identity checks for new users (Reg LOA=0). But they are suffering from the weakness of password authentication and would like to have something stronger (AuthN LOA>=NIST[5] level2?); every now and then a kid manages to hack his friend's password, and is then able to do nasty things in the virtual world.
LoA Details:	This is an interesting scenario since the virtual world is presumably all that exists (ignoring the fact that you might wish to prosecute someone in real life for stealing virtual resources – which has happened). Therefore the self asserted attributes could be said to have a very high Registration LoA because the creator of the virtual character gave it its virtual attributes. Its akin to God being the registration authority. In this case the Session LoA equates to the Authn LoA, which is initially low. Improving the authentication mechanism will increase the Authn LoA and Session LoA, since we are assuming the Registration LoA is at the maximum value possible. This scenario would change once we had virtual IdPs and virtual AAs (such as virtual universities) that could hand out virtual attributes (such as virtual degrees). Then the virtual world would mirror the real world again.

Case:	UC5: Parental Access
Scenario:	Using Government issued IDs to allow parents to access resources at their children's schools.
LoA Details:	Government issued IDs will presumably have high registration requirements involving links to existing Government records. The Registration LoA will therefore be high. Various different authentication technologies are envisaged, including username/password and smart cards, each of which will have an associated Authn LoA. There may then be a further registration process to establish the link between the parent and the child; this will often be performed by the school or Local Authority and, at least if smart cards are used, may provide the lowest component of the LoA.

Case:	UC6: Access to Grid Resources
Scenario:	A researcher wishes to use a range of Grid resources; the providers of these resources each require a high Level of Assurance that the researcher is a valid user (typically, has agreed to certain terms and conditions).
LoA Details:	Both technologies and processes need to be in place at both Registration and Authentication to ensure the session LoA is high (eg NIST level3).

Case:	UC7: Student Identity Received from a Trusted 3rd Party
Scenario:	Rather than undertake (costly) face-face checks with all incoming students, a University may rely on the evidence of identity provided by a 3rd party such as a school or central university clearing service; continued use of this identity information by the university reinforces its validity over time. SPs wish to know that the university "accepts responsibility" for the identities of users from that university.
LoA Details:	The Registration LoA in this scenario is fairly low, lower than NIST level2, but is higher than NIST level1 (self-registration, such as provided by OpenID[2]). In most cases, the AuthN LoA would be at NIST level2 (username and password), but the overall LoA is the lower of the two. Many SPs are willing to accept a level of identity assurance lower than 2, but would want it to be higher than 1, so provision needs to be made to allow an intermediate (level 1.5??) registration LoA. In Australia, this LoA is called the "floor of trust".

Case:	UC8: Mobile Phone Usage
Scenario:	
LoA Details:	

Definitions

1. Credential Validation: the process of validating that a particular credential is both authentic (i.e. not tampered with or revoked since issuance) and trusted (i.e. issued by a trusted authority);
2. Identity Verification: the process of linking a real-world person to an account;
3. Verification of Attributes: the process of linking real-world attributes to an account;
4. The ceiling of the LOA: meaning the level of assurance above which performing either identity or attribute verification is not worth because it would be easier for an imposter to go and forge the credential (or beat up the rightful owner). For example, the ceiling of LoA is much lower for an 8-bit PGP key than for a 4096-bit one, for example
5. Floor of trust: [in Australia] the basic LoA available within the Federation – higher than NIST level1 (self-registration), but usually lower than NIST level2 (face-face identity verification at registration).
6. Level of handling of credentials: no matter how high-quality an authentication or authorization token, the strength can be reduced by user bad practice e.g. lending someone your credit card and PIN.

References

[1] FEIDE:
http://feide.no/

[2] OpenId:
http://openid.net/

[3]Shintau project:
http://sec.cs.kent.ac.uk/shintau/

[4] Liberty Alliance:
http://www.projectliberty.org/

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