Trusted third party based ID federation, enhancing privacy and lowering the bar for connecting

Author:
David Simonsen, WAYF - Where Are You From
Jacob-Steen Madsen, WAYF - Where Are You From

The Agency for Library and Media, H.C. Andersens Boulevard 2, DK-1553 Copenhagen V, Denmark, david@wayf.dk, jsm@wayf.dk

Keywords: Identity federation, trusted third party, privacy, protocol translation, consent, SAML2, Shibboleth, simpleSAMLphp, (in)visibility,

The Danish ID federation for education and research, WAYF - Where Are You From, is combining the two existing ID-federation models:

1) the Shibboleth model based on decentral login-systems at the connected institutions and a full mesh (many-to-many relation) between home institutions and service providers (implemented in i.e. UK, Switzerland, Finland etc.)(see Fig. 1) and

2) the hub-and-spoke model where institutions connect to a central login-service (one-to-many relation) while keeping user data and identity management at the institutional level (as used in FEIDE, Norway) (see Fig. 2).

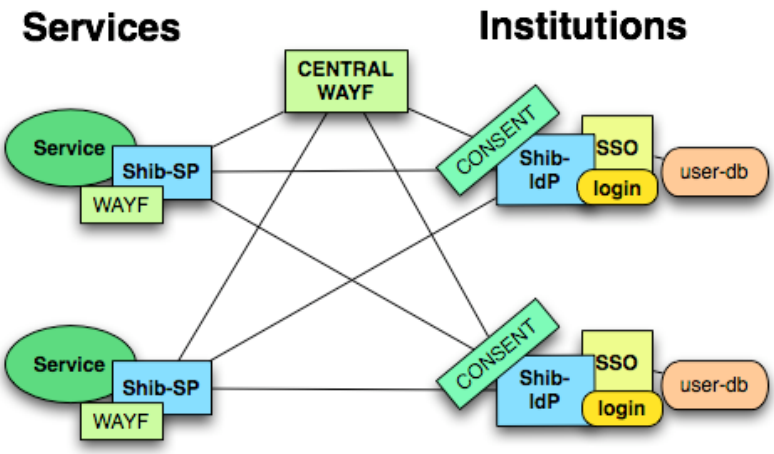

Fig. 1. Shibboleth architecture: full mesh, decentral login and decentral identity management

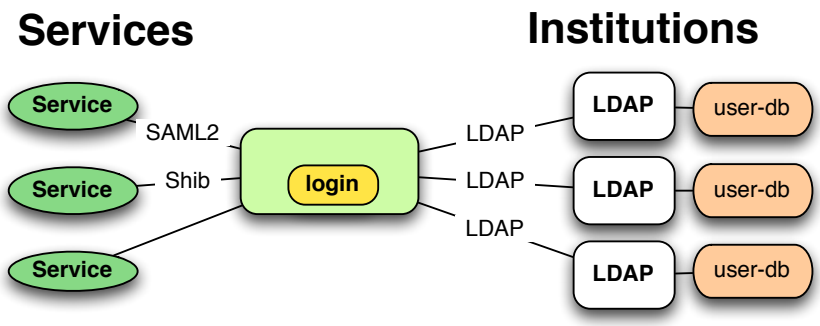

Fig. 2. Hub-and-spoke architecture, central login, decentral identity management

It is clear that in large complex systems decentralization is the only viable model. Centralized models, like the one described here, will at a certain point no longer scale. But for countries like Denmark, where many institutions (and their IT-departments) as well as the involved communities, are small the establishment of a trusted third party may be a sustainable solution.

WAYF has spent several years building an ID-federation for education and research. First a classic Shibboleth model was introduced but proved to require too much specialized know-how both for the institutions and for the service providers.

Then the hub-and-spoke model with centralized authentication was suggested.

This in turn conflicted with the already established single-sign-on systems and the substantial branding efforts already running. The resulting trusted third party model, TTP, was a combination of the two: a hub-and-spoke model with decentral authentication where the trusted third party acts as a single proxy-identity provider towards the connected web based services - on behalf of all connected institutions.

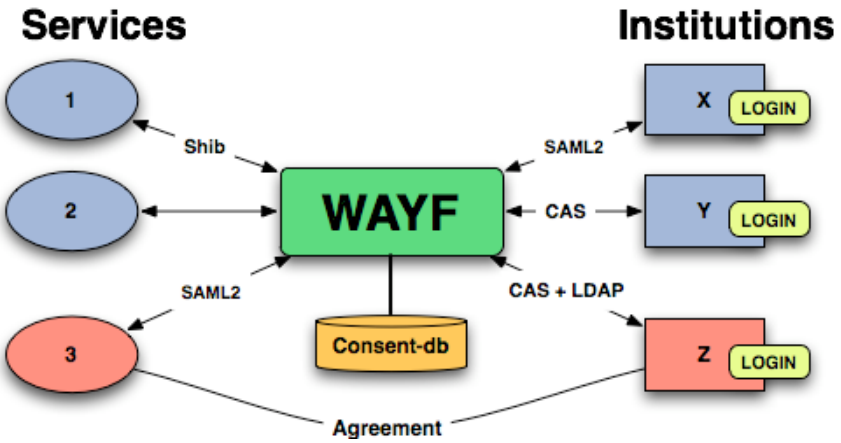

Fig. 3. Hub-and-spoke architecture, decentral login, decentral identity management

While implementing the technical solution, several new ideas about how to run an ID-federation emerged. The functional and architectural benefits of the combined model now seems to be several:

- Users are branding wise 'kept close' to the institutions as they always authenticate at the home institution, a cardinal requirement for for the institutions

- Both institutions and service providers only have to establish a single connection in order to gain access to all members of the federation (one-to-many relation)

- Most institutions already have single-sign-on systems (SSO) which can now be extended with the total number of services connected to the federation

- As the trusted third party has the role of proxy identity provider

- Several protocols for connecting institutions are supported to let institutions concentrate on identity management in stead of protocols (that may change over time)

- Complex tasks like collecting and managing users' consents to attribute release are centrally performed by a specialized entity

- Certain attributes (like organization name) are provided by the trusted third party, so that no one can claim to be someone else

- The attribute eduPersonTargetedID is centrally calculated (a hash value) and can therefore not be deciphered by neither the service provider nor the institution nor the two in combination. All three parties have to corporate - which makes the task non-trivial and therefore unlikely. Thus the users' privacy is better protected than in the two other models, as the personal pseudonym stays anonymous to both service providers, institutions and the trusted third party until an important matter brings the three parties together to identify the user.

- The wayf-function (where are you from) can either be integrated in the web based service or provided by the trusted third party. Even when integrated in the service, the centrally provided consent functions are working.

- The attribute release policy (ARP) is negotiated centrally with the trusted third party when connecting to the federation. The ARP must support the principle of proportionality as defined by the EU directive on transfer of personal information. With the third-party-model all users at all institutions consent to releasing the same amount of information to a given service - hence the institutions do not have to negotiate with individual services.

- Institutions may choose not to release users attributes to a given service (opt-out), but the general principle is that as a starting point all services can receive attributes from all institutions

- The users' attributes are not stored centrally by the trusted third party, but information is kept for the length of the browser session in order to provide the single-sign-on functionality (for example a WAYF-session is 8 hours)

- The trusted third party's legal status is 'data processor' for the connected institutions. This places legal responsibility for the users' personal information with the institutions and consequently lowers the audit requirements for the trusted third party. A regular service contract is signed by the connected service providers.

- When ID-federations connect, or confederate - both nationally and internationally, neither the institutions nor the service providers have to adjust their technical setup since the new connections are provided centrally by the trusted third party. When such new structures are established new partnerships can easily be established as a result of the now accessible services and all the users at the interconnected federations' institutions.

Room for improvement

All this being said, the third-party-model, as implemented by WAYF is still in its' infancy. Only a small albeit growing number of users are using the infrastructure and questions about usability are brought to attention already now. The general model for how users are redirected to authenticate and the following data flow is summarized below (Fig. 4).

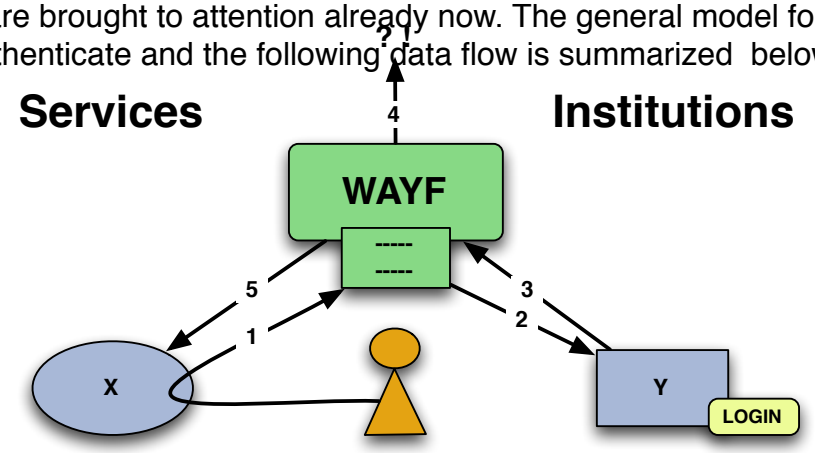

Fig. 4. User experience and data flow diagram

1. The web based service's login function sends the user to the WAYF web based source where the user chooses his/her institution.

2. If the user is not already logged in he/she is redirected to the institution's login page (he/she will automatically transferred back to the web based service).

3. After login at the institution information about the user is sent to WAYF.

4. The WAYF web page presents the information that will be forwarded to the web based service. The user gives his/her consent with a click on a button. The user can indicate with a 'tick' that the consent should be remembered when visiting the same web based service in the future.

5. WAYF sends the information to the web based service. If the web based service can approve the user based on the information then access is granted.

Already now questions about both the overall usability as well as more specific use cases arise:

- If the central wayf-function is used, what reactions should be expected if a service does not authorize users from institutions that can be chosen in the list of connected institutions? There seems to be no simple way of warning the user that he will not be let in.

- If the central wayf-function is used, what reactions should be expected if an institution does not let their users access a given service? One solution could be to 'gray out' institutions that the trusted third party knows do not allow attributes to be sent to a given service.

- can the users manage the 'two step' model where the

- How does the trusted third party earn to be trusted by the individual user? Why should they trust it? Is it up to the connected institutions to educate the users?

- How visible should the trusted third party be? Should it be invisible to the user?

Solutions for the above mentioned questions, and certainly many more in the near future, must be provided. A new use case, support for age verification, has been requested recently and calls for yet another are of close corporation with the international ID-federation community.

Vitae

David Simonsen has been involved in work with e-IDs and ID-federations for education and research since 2004, developing and deploying the WAYF federation since 2005. Before that he was co-charing the TERENA task force mobility which developed 'eduroam'.