Implementing Shibboleth at a UK National Academic Data Centre

Ross MacIntyre
MIMAS – Manchester Computing, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
ross.macintyre@manchester.ac.uk

David Chaplin
MIMAS – Manchester Computing, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
david.chaplin@manchester.ac.uk

ABSTRACT

The UK education sector is embarking upon the adoption of Internet2's *Shibboleth* software for federated access management. This paper recounts the early experiences of a large academic data centre in implementing support for Shibboleth across its range of services. It covers the practical approach adopted, a worked example and the significant issues raised. Familiarity with federated access and identity management is assumed.

KEYWORDS

Shibboleth, Authentication, Authorisation, Federation, Federated Access, Federated Identity.

1. INTRODUCTION

The University of Manchester provides a range of both local and national services through Manchester Computing. It is responsible for the provision and support of computing services to the University as well as to members of other academic institutions throughout the UK, Europe and beyond. A major national provision is via MIMAS (Manchester Information and Associated Services) [9]. MIMAS receives government funding, via the Joint Information Systems Committee (JISC) and Economic and Social Research Council (ESRC), to act as a National Data Centre providing the UK Higher Education, Further Education and Research communities with networked access to key data and information resources.

MIMAS currently hosts a heterogeneous range of services: the UK Census; Socio-Economic data; Satellite and Digital Map data; bibliographic resources, including the UK JSTOR mirror and the 'ISI Web of Knowledge Service for UK Education'. A variety of authentication, authorisation and registration requirements exist across the portfolio. These vary from freely available resources, with anonymous access, to restricted IP access, right on up to personal usernames and passwords and individual registration, i.e. identification. (A fuller contextual description of the UK environment is available elsewhere [4].)

JISC has devoted a significant part of its development funding to access management issues, including funding MIMAS [10] to implement support for Federated Access, using *Shibboleth* [15], across all its services - if possible. Note that some services are run remotely and MIMAS may have no direct involvement in the access control method implemented. Shibboleth is described briefly in the next section and the paper describes progress to date.

2. IMPLEMENTATION

Internet2's Middleware Architecture Committee for Education (MACE) lead the development of the Shibboleth software, which in their own words: "leverages campus identity and access management infrastructures to authenticate individuals and then sends information about them to the resource site, enabling the resource provider to make an informed authorization decision.

For example, when a student requests access to a protected video clip, her home organization requests her to authenticate (if she has not done so already) and then passes on the information that she is enrolled in Biology 562 to the site housing the video. The video provider uses the fact that she is enrolled in this course to determine her eligibility to access the video."

Shibboleth does not carry out authentication itself. Instead Shibboleth defines a set of protocols for the secure passing of identity information between institutions and service providers. It relies on the institution to establish identity, and on the service provider to confirm access rights, given information about institutional affiliation and possibly other user attributes. It is written in SAML (Security Assertion Markup Language), an international standard developed by the OASIS Security Services Technical Committee. Shibboleth has defined a standard set of attributes; the first set is based on Educause's eduPerson [2] object class that includes widely-used person attributes in higher education.

How authentication is carried out by the institution, and how rights management is carried out by the service provider is left up to the respective parties. In so doing, Shibboleth depends on a certain level of trust. Service providers need to be confident that the institution or organisation that the user belongs to has a robust and up-to-date authentication system in place.

This need for trust leads to the concept of federations. Federations are groups of similar organisation such as universities who have agreed to a common set of policies. They are typically being established at a national level. For example US higher education has established a federation known as InCommon[6]. The equivalent in Switzerland is known as SWITCHaai [17] and in Finland as HAKA[5]. In the course of 2005 the JISC plans to establish a federation covering the UK's higher and further education sectors[13], though a development federation, SDSS[14], already exists, administered by the other UK academic data centre, EDINA, based in Edinburgh.

2.1 Existing National Access Management System

As mentioned in the Introduction, MIMAS resources are subject to a variety of access policies and methods. However, the majority are subject to authentication and authorisation via 'Athens'[13]. In August 2000, Eduserv were contracted by the JISC for the Provision of Authentication Services to the UK higher and further education community, their solution is called Athens. Hence the community has an existing co-ordinated approach already in place, which is both a help and a hinderance. It helps that lines of communication and responsibilities are very well established and a consistent understanding of access control exists amongst the institutional administrators and management. This means the benefits of the federated access model associated with Shibboleth are easier to argue, but its adoption requires significant adjustment to well-established, working procedures. Furthermore, the end-user is not (usually) interested in the access control methodologies employed and resents any changes necessitated as an irritating interference in their work patterns.

Note that Eduserv have created two federations themselves, one development and one production and are including support for Shibboleth and SAML protocols in the Athens code.

2.2 Approach Adopted

A survey was performed across all services and their constituent applications and the following key aspects identified:

- Was Registration required?
- What Authentication method(s) were in place?
- What Authorisation method(s) were in place?
- Which Attributes were (currently) required by the service?

In discussion with service support staff, a phased approach was proposed for implementation over two years, summarized below:

Phase 0 - No work required as the resources are freely accessible.
Phase 1 [April-May 2004] - Initial testing of Shibboleth(v1.2) code.

- A dumb test page ("Hello World"), purely internal to MIMAS, though access to this page would be tested from two other sites that have Shibboleth origin code installed.

Phase 2 [June 2004-May 2005] - MIMAS services, first group, comprising:

- A bibliographic reference service (Zetoc [18]), though this includes a current awareness alerting facility, which requires a unique identifier to be passed in order to identify the correct, stored alert lists. This being an (uncomplicated) example of a service requiring a user attribute;
- Map data, where MIMAS control authorisation and require registration in order to create a user account;
- UK Census data, which requires collaboration with and is dependent upon UK Data Archive's Census Registration Service (CRS) [1];
- Mirror of the e-Journal archive JSTOR, requires liaison with US JSTOR team.

Phase 3 [June 2005-May 2006] - MIMAS services, second group, comprising:

- Online teaching materials;
- Satellite Datasets;
- Socio-economic data, which follows naturally from the Census work as it also relies upon CRS;
- A commercially sensitive web site containing consortial pricing information.

Phase 4 - MIMAS services where reliant upon external suppliers. Including:

- A&I database;
- Chemical information database;
- eJournal Discovery Tool;
- Learning Materials Repository.

Commitment and plans were to be requested from the supplier.

2.3 Progress

The proposed approach was accepted and a formal project, ShiMMeR (Shibbolising MIMAS eResouces)[16], was initiated in 2004 and work began immediately. MIMAS formally joined the InQueue[17] and SDSS federations. The former is purely for testing and is inherently untrustworthy. The latter, though not full production strength, is much more stringent and requires the use of digital certification (e.g. GlobalSign[4]) to formally establish trust.

Implementation has proceeded on schedule, though some reordering has occurred. Three of the four Phase 2 services have been 'Shibbolised', together with the Satellite Datasets, brought forward from Phase 3. The UK Census data has been moved to Phase 4, though discussions have taken place with the UK Data Archive on dependencies.

2.4 Example Implementation

Included below is a real example, illustrating the effect of the implementation of Shibboleth and the consequential separation of authentication and authorisation in the Map data service 'Landmap' [8].

2.4.1 Existing – Athens

On the Landmap website there are a number of paths that result in an authentication challenge. Just one case is considered here, the request to "Download Landmap Data" by an unregistered user from a subscribed institution:

- Selection

After selecting some data to download and submitting the request (at which point the data is copied into a temporary download directory), the user is prompted to "Click here to authenticate and then download your data".

- Authentication

The Athens Authentication Point appears and the user is prompted to enter their Athens credentials (username and password). Assuming this authentication is successful (which the site remembers by sending cookies to the user's browser).

- Authorisation

Athens then checks whether the user is authorised to use the Landmap download services. So note that authentication and authorisation are effectively bundled into a single check. Following the granting of access (remember that the user has successfully authenticated and is from a subscribed institution) a check is performed to ascertain whether the user has previously registered with the site.

- Registration

In our scenario the answer is not and so the user is prompted to enter registration details - a prerequisite for using the Landmap download services. These details (surname, first name, email address and department), which are not validated at present, are stored for future reference and the user is informed that the download can proceed.

- Download

In the download procedure the data is copied from the temporary download directory to the user's browser, the details of the download are logged for auditing purposes, and the temporary directory is then deleted.

2.4.2 New - Shibboleth

A Shibboleth protected site has been created and is running side by side with the existing Athens protected site on the same MIMAS service machine, but on a separate web server.

The site has been signed up as a service provider in the SDSS federation and uses a certificate signed by the SDSS CA (note that browsers will alert users of a trust issue with this certificate unless the SDSS CA certificate is manually installed in the browser).

The typical workings of this Shibbolised site are now outlined by considering the same use case detailed previously for Athens, that of a request to "Download Landmap Data" by an unregistered user from a subscribed institution:

- Selection

After selecting some data to download and submitting the request (at which point the data is copied into a temporary download directory), the user is prompted to "Initiate Download". This calls a script that resides in a Shibboleth protected secure area on the web server.

- Authentication

Assuming that the browser has just been opened, the user is directed to the WAYF ("Where Are You From?") service provided by the federation. This offers a list of all identity providers signed up to the federation, from which the user selects their home institution and enters their local credentials. So the authentication procedure is now the responsibility of the individual institution.

- Registration

If this authentication is successful, a Shibboleth session is established, and the user's home institution releases a number of attributes about the user in response to a request from the Shibbolised Landmap site. Note that only the *eduPersonPrincipalName* attribute is required by the current site, and that this *eduPersonPrincipalName* attribute is in the form of an email address comprising a username and the domain of an institution/department, e.g. jbloggs@institution.ac.uk.

Assuming this attribute is received by the site (the user is not permitted to proceed otherwise) a check is performed to ascertain whether the user has previously registered. In this scenario the answer is not, and so the user is prompted to enter registration details comprising three mandatory, though not validated, fields (first name, surname and department). These details are stored for future reference, along with the *eduPersonPrincipalName* value, and an authorisation procedure is initiated. It is intended to automate this registration step using appropriate attributes in the near future.

- Authorisation

This process checks the domain name obtained from the *eduPersonPrincipalName* attribute against a list of subscribed institutions maintained by the Shibbolised Landmap site, and grants authority to proceed only if a match is found. So authorisation is now performed in-house by the service provider. In the present scenario the user's institution has subscribed to the requested service and the user is informed that the download can proceed.

- Download

The download procedure is as before: the data is copied from the temporary download directory to the user's browser, the details of the download are logged for auditing purposes, and the temporary directory is then deleted.

3. CONCLUSION

The implementation of Shibboleth itself is technically straight-forward, however the management and user-interface issues are not. The policy framework that needs to exist is evolving, but has potentially profound effects upon implementation.

Below are some substantial questions that arose during the first stage of the project, but which remain unresolved.

- Which federation?

At present, those associated with taking Shibboleth forward appear to be fairly casual in their approach to federation establishment and membership. However, this must change, as the federation will establish 'terms and conditions' that must be adhered to by members. There are significant user-interface issues associated with establishing which federation a user and their institution belong to. For example, federations typically have an associated WAYF service for determining institutional membership, though there may be many. Notice also that support for membership of multiple federations was only introduced in Shibboleth version 1.3, the first beta release of which was made available for download in June 2005.

Quite what approach global resource providers and rights holders adopt regarding federation membership or participation, is certainly intriguing.

- What attributes can/should be requested/required/recorded?

By way of example, one MIMAS service (Zetoc) needed to know a user's affiliation, since the service includes institutional preferences. However, also having some persistent identifier would have made the sessioning more straight-forward and resilient. Add to this that a persistent personal identifier was needed for the separately authorised personal alerting function (to link to previously declared search terms &/or journal lists for notification). Is it not reasonable to demand both an institutional identifier (*eduPersonScopedAffiliation*) plus persistent identifier (*eduPersonTargetedId*) from the outset?

More generally, where should decisions be made about what attributes are reasonable? For the JISC community and more widely? Would it be a subtask of Federation policy management? This could be seen as the flip side of Attribute Release Policy (ARP) - referred to as Attribute Demand Policy (ADP), perhaps[12]?

How will commercial concerns compromise between what is required for the application/service to work as intended, what the rights holder may require and what their Marketing Department may like?

Where should these ADPs be recorded? In the JISC Information Environment context, this could be in extensions to a JISC resources registry. More widely, encouraging the establishment of ADP registries of the attributes that vendors typically demand to allow access to each of their identifiable services/resources (via an institutional license) would help to establish the international requirements for *eduPerson*.

- Replacement for local/distributed registration data?

During implementation, especially where non-trivial registration is required, how does one deal with the situation where only some attributes can be provided by the institution?

- User interface(s) during process?

In all this, the end-user just wants to get to what they want and are entitled to access. They do not wish to be interrupted by middleware implementation issues. However, in order to present the right information to the end-user and ask them the right question, the system almost needs to know the answer beforehand.

ACKNOWLEDGEMENT

The ShiMMeR Project is funded by the Joint Information Systems Committee (JISC) of the UK Higher and Further Education Funding Councils. The authors wish to acknowledge the assistance of colleagues: Kamie Kitmitto and Andrew Weeks, MIMAS; Sandy Shaw and Fiona Culloch, EDINA; John Paschoud and Simon McLeish, LSE.

REFERENCES

1. Census Registration Service (CRS): http://census.data-archive.ac.uk/

2. EDUCAUSE *eduPerson* Object Class: http://www.educause.edu/eduPersonObjectClass/949
3. Eduserv Athens Service: http://www.athensams.net/
4. GlobalSign: http://www.globalsign.net/
5. HAKA Federation: http://www.csc.fi/suomi/funet/middleware/english/HAKA_final_report.pdf
6. InCommon Federation: http://www.incommonfederation.org/
7. InQueue Federation: http://inqueue.internet2.edu
8. Landmap: http://www.landmap.ac.uk
9. MIMAS: http://www.mimas.ac.uk/
10. Morrow T., Borda A. and Robiette A., Shibboleth, Connecting People and Resources, *JISC Briefing Paper*,
11. http://www.jisc.ac.uk/uploaded_documents/JISC-BP-Shibboleth-v1-final.pdf
12. Paschoud J., Shibboleth and SAML: At last, a viable global standard for resource access management, *New Review of Information Networking*, Vol.10, No.2, 2004, pp 147-160.
13. Paschoud J., "Re: Reed puts 300,000 at risk of online fraud: A good argument for using Shibboleth? Attribute Demand Policies", an email message sent to the e-mail discussion list: jisc-shibboleth@jiscmail.ac.uk 14/04/2005.
14. Robiette A. and Morrow T., Blueprint for a JISC Production Federation, *JISC Position Paper*, http://www.jisc.ac.uk/uploaded_documents/JISC_Fed_doc_full.doc
15. SDSS: http://www.sdss.ac.uk
16. Shibboleth: http://shibboleth.internet2.edu/
17. ShiMMeR Project: http://www.mimas.ac.uk/shibboleth/
18. SWITCHaai Federation: http://www.switch.ch/aai/
19. Zetoc: http://zetoc.mimas.ac.uk