Distributed authentication for the Western Australian University Libraries

Peter Green
e-Library Development Librarian
Curtin University of Technology
P.Green@Curtin.edu.au

Abstract:
The need to establish a means of achieving reliable, automated inter-institutional authentication was identified by the Western Australian Group of University Librarians (WAGUL) in 1999 as a strategic priority and a project was subsequently established as a result of a substantial grant from the Commonwealth Development Pool (CDP). The WAGUL Authentication Project (WALAP) first reported on its progress at VALA 2002 and now follows up with a report on the achievements to date and the challenges to come.

Introduction

Start at the beginning

The need to establish a means of achieving reliable, automated inter-institutional authentication was identified by the Western Australian Group of University Librarians (WAGUL) in 1999 as a strategic priority and a project was subsequently established as a result of a substantial grant from the Commonwealth Development Pool (CDP). A paper was presented at VALA 2002. At that time the project team had successfully scoped the project, written and advertised a substantial Request for Proposal (RFP) and was at the post-RFP stage with a preferred supplier having been selected and protracted contract negations underway. However much work remained to be done by the project team before the envisaged infrastructure could be delivered. Now, much further down the track, a report is made on the progress, successes and limitations of the distributed authentication infrastructure that has been constructed. Particular attention will be paid, in this paper, to the architecture implemented, the constraints imposed by attempting authentication and authorisation across institutional boundaries and the difficulties of leveraging the infrastructure to the advantage of the libraries and the universities.

Synopsis of the Plot

Jump to the last page

But what does it do? The objective of the project can be stated fairly simply "to assist in the development of an authentication infrastructure to allow WAGUL libraries to provide access to online resources including scholarly information in electronic formats" (Green 2002b). However a few theoretical use cases might suffice for this paper to paint a picture of what the distributed authentication infrastructure enables.

A student from Curtin University of Technology (Curtin) sits down at a workstation in the library at Murdoch University (Murdoch) to access a web resource that is shared by the universities (in some consortium arrangement). The Curtin student enters the student number and password that she uses at her own university, and indicates that she is a Curtin student. The web application is authorised to interrogate the directory located at Murdoch and does so, but the directory at Murdoch realises that it doesn't hold the information for this person. However the Murdoch directory is authorised to interrogate the Curtin directory and does so. The result is passed back to the web application which allows the student access to the resource. The web application is hosted locally at Murdoch and has no direct access to any directory other than that at Murdoch. The administration of the web application is done locally at Murdoch.

A student from Edith Cowan University (ECU) sits down at a machine on their home campus to access a Learning Management System (LMS) located at the University of Western Australia (UWA). The student is enrolled at ECU but doing a unit serviced by UWA. The student logs into the LMS using their ECU student number and password. The LMS is authorised to interrogate the UWA directory, which in turn interrogates the ECU directory. In this case the student must authenticate correctly and also be enrolled in the particular unit. Both of these are true and the LMS receives confirmation of the authentication and authorises access to the unit based on information provided by the ECU directory. The LMS is hosted

and managed at UWA but the information on which it bases its authorised access is managed by ECU.

The advantage of the distributed authentication is that information about a student need only be managed at one place, their home institution, yet becomes available in the distributed context to any applications that are authorised. This is the essence of an underlying authentication infrastructure, taking the hard work out of authentication and reducing it to a plug and play solution. To create infrastructure is difficult and time consuming, but this then becomes work that doesn't have to be re-created by those who require authentication.

The Story Continues

Where we left off

The paper that was previously presented to VALA on the WAGUL Authentication Project (Green 2002) was written in the latter half of 2001. At that time the project had been scoped, a Request for Proposal had been written, advertised and responses evaluated. A recommended supplier had been nominated and contract negotiations were about to begin. By the time the paper was presented to VALA in February 2002 the contract negotiations were still underway and work hadn't begun. It had taken a good part of the first year of the project to reach this point. The nature of the project, involving five universities, and the nature of the contract where it was not fixed cost and where contractors would be on site at all five campuses during the implementation, meant some delay in negotiations between the five universities before the final contracts could be signed with the supplier. The extended time required for such inter-university negotiations has been demonstrated more than once during the life of the project and should serve as a warning to others not to under estimate the time required, even given good will from all parties in the process.

However, once the legal paperwork was completed the substantial work of creating the distributed authentication infrastructure could begin. This was a period of intense activity involving many staff from the five universities in design workshops, testing, training and technical implementation. The structure of the directories was designed, prototyped and then rolled out sequentially to all five sites. The final installation was tested using a prototype web application, 100,000 dummy students at each site and passed all testing requirements with flying colours. The contract with the supplier was brought to a close in late 2002, within the original budget estimates and very close to the original time estimates. This was a notable achievement given the uncertainties associated with a time and materials contract and the difficulties of coordinating the work over five sites. To reach this point had taken the better part of two years. However that wasn't the end of the story, but merely the first step. With these foundations laid the real work could begin.

Fleshing out the Characters

A full cast is required to start

Before continuing the story, a moment to flesh out some of the detail of what has been constructed so far.

The technical base upon which the distributed authentication infrastructure is built is described in more detail elsewhere (Green & Reid 2003) but in brief each of the five universities hosts a directory (a type of database) that contains information about staff and

students that can be 'looked up' by an application. The interesting part is the ability for the directories themselves to look up information held at the other universities and provide an answer back to the original application, without the need for human intervention. This is the 'distributed' part of the infrastructure, and what gives this project some claim to innovation. The administration and ownership is not distributed and resides solely with each university. The permission to conduct such authentication across institutional boundaries is pre-defined and pre-agreed, thus allowed for a fully automated solution.

For a person to be authenticated that person must be in the directory and have attributes. This is beyond dispute. To authenticate requires an identifier and a password. This is a non-trivial task to establish and maintain, but is relatively straightforward in concept.

The construction of the schema for the directories is covered in some detail elsewhere (Green & Reid 2003), but suffice it to say that a new schema was developed for this purpose and given the name auEduPerson. To meet the authentication requirements WALAP agreed on two attributes to meet the usual username/password form of authentication, *auEduPersonID* and *userPassword*.

How they would be populated was a decision for each site. Provided that they were internally unique, and known to staff and students, their format is not relevant. Generally, however, student number and staff Identifier (ID) were chosen as they are already internally unique and in use for authentication. An attribute was also included for certificates, *userCertificate*, but these are not yet in common use and the attribute was provided for future use. A Certificate is a unique, international, complex and secure identifier/number that is proposed as the basis for identity management in the future.

Now we get to the muddy ground. Authorisation is based on knowledge about the person. One would imagine, unless one knew better, that it would at least be easy to decide if a person is a staff member (or a student) of the university. How these terms are understood and defined however varies from university to university, and they are sometimes poorly defined with a university with different areas having different practices. An agreement was required to 'define' possible values for these attributes without attempting the impossible. A minimal approach was taken with the expectation that time and usage would better inform the decision making process. A common theme of the definitions is trust. The onus for deciding how a person was defined was left to each university.

To provide a classification that could be broadly used in the inter-institutional context two attributes were created. The first, *auEduPersonType,* would allow for a broad distinction between staff and student, allowing for the possibility of an identity that was not considered either but might be in the directory because some other relationship with the university, for instance a visiting speaker. However a person could only be just one of these within the directory and the attribute was defined as mutually exclusive. For a person who is both staff and student it was agreed that the common practice was already to treat them as two separate 'identities' and that a person would be one or the other at any one time depending on which ID they chose to use.

To then allow for finer grained authorisation the *auEduPersonSubType* attribute was created. This would allow an application to decide not only that a person was a student but also what sort of student. For instance if an application wanted to provide access to all staff and postgraduate students a simple rule could be written based on an entry having 'staff' in

auEduPersonType **or** having both ‘student’ in *auEduPersonType* **and** ‘postgraduate’ in *auEduPersonSubType*. A small number of possible values were agreed for *auEduPersonSubType* but time and usage would better inform the usefulness of these values. This attribute is multi-valued. For instance a student could be both ‘postgraduate’ and ‘external’, both of which would be useful values in different contexts.

The other major area of authorisation relates to units. There are probably many ways of cutting this, but for our structure we created the *auEduPersonActiveUnit* attribute to store the unit in which a person is involved. The value of this attribute is based on each site having a unique code for units and a local definition of ‘active’. The attribute is multi-valued allowing students to be enrolled in multiple units.

There may be instances when a person is retained in the directory for some purpose, but would be considered ‘expired’ and so the *auEduPersonExpiryDate* attribute was created for this purpose. However it is unclear as to the use of this in practise and it was agreed that a null value would indicate ‘current’ under the agreement.

These six attributes form the basis on which authentication and authorisation will be conducted. In the process of defining the object classes for the directories other attributes were originally included. These remain within the definition, but after some reflection the project team decided that some of these attributes held information that was more sensitive and less clearly justifiable for authorisation and thus they were not eventually included in the initial sharing agreement. In the fullness of time the need for some of these may become clearer and the number of ‘active’ attributes may grow.

The ability to describe persons using a limited number of attributes, and having the technology in place to store and interrogate the information, means that only one step further is required before the distributed authentication is in place. However this extra step has proved to be more difficult than originally expected. The extra step is to populate the directories with staff and student information, accurately and in a timely manner.

It might be supposed that the information required is already available in one or other university databases and that the means of automating its extraction and transfer into the directory also exists. Now this is largely true, but the logistics of creating the pipelines was left as an exercise to the technical staff of each university after the contracted implementation was completed. In some respects this was inevitable, as the required local knowledge was already in place and the task wasn’t well suited to outside contracting. Populating the directories with even this small number of attributes was not as simple as might be supposed and required some substantial work to achieve. This would prove to be a slow grind for the project. At the time of writing 3 of the 5 directories are populated in an ongoing way, and two are in process with expectations of completion by November 2003. This has taken a lot longer than anticipated and highlights the difficulty of dealing with real data and real systems. It also highlights the lack of excess capacity within the universities to conduct ‘additional’ work even when project funding is available.

The Plot Thickens

A matter of trust

Did I say only one extra step was required once the attributes were agreed and the technology was in place? This is not quite true. Even before the directory populating was begun a missing link had been identified. This was the question of trust.

The distributed authentication infrastructure, while not requiring that any sensitive information about staff or students be stored outside of the home university, does allow some limited access to that information. This enters difficult territory and underpins the tension between collaboration and competition, with the added spice of litigation. It was decided that a formal agreement was required to state the purpose for which access to the directories was being allowed, the use to which that access could be put and some mutual indemnity in the case of problems. The first version of the Mutual Confidentiality Agreement was drafted towards the end of 2002 and the signature of the fifth vice-Chancellor was obtained in the second half of 2003. This was not because the agreement was contentious, but because the logistics of consultation and agreement between five parties of this nature is very time consuming, as has been demonstrated a number of times by the project. However this agreement puts in place the final piece of the puzzle that will allow the distributed authentication infrastructure to be used.

Back to the Beginning

And in the next episode…

What about the libraries? The distributed authentication infrastructure was consciously scoped to be a university level solution even though the project was being led by the university libraries. This is one of the consequences of the electronic age; it is difficult to work effectively and economically as an island even within a university. However, just because the distributed authentication infrastructure was created with full involvement at the university Information Technology (IT) level and the shape of the directories and the attributes and the legal agreements were all done to allow functionality at a university level, this doesn't mean that the project didn't continue to be driven by the libraries. In the great effort to get the distributed authentication infrastructure in place they didn't neglect to look towards the next step and the production use of the distributed authentication infrastructure. The project may be complete once the infrastructure becomes fully available but the work will be ongoing to take advantage of the opportunity that it presents.

One good example of this ongoing work is the integration of LIDDAS into the distributed authentication infrastructure. "LIDDAS (Local Interlending and Document Delivery Administration System) is a comprehensive automated interlibrary loans management system developed to enhance both the requesting and delivery of items to researchers" (Bronleigh 2003). This major WAGUL and Australian Vice-Chancellors' Committee (AVCC) project was always seen as the prime candidate to be first cab off the rank. LIDDAS has also been in gestation for an extended period and will be familiar to a library audience outside of Western Australia. The WAGUL implementation involves four of the five universities with the service being hosted at UWA on behalf of the consortium. Amongst other complexities, authentication has been an issue and it is expected that the distributed authentication infrastructure will provide exactly the solution that is required. At the time of writing, early December 2003, the latest version of the VDX software (Virtual Document eXchange from

Fretwell-downing upon which LIDDAS is built) was being installed, a version with LDAP capabilities, and a subsequent release was being awaited to provide the full functionality required to use the distributed authentication infrastructure in a fully secured manner. LDAP (Lightweight Directory Access Protocol) is the standard protocol used for accessing directories and is increasingly being included as base functionality by product developers to provide authentication services. The integration work for VDX is being conducted by UWA on behalf of the project and is seen as a good demonstrator of the distributed authentication infrastructure fulfilling its purpose.

Three of the universities (UWA, Murdoch and ECU) use the same Integrated Library Management System (ILMS), Innopac from Innovative Interfaces. With UWA taking the lead each of the three entered into a Development Partnership with Innovative Enterprises to develop the LDAP functionality that would allow Innopac to take advantage of the distributed authentication infrastructure. While this functionality would have initial benefits to each library, the way would then be open for cooperation between the libraries to be further development taking advantage of the distributed authentication infrastructure. Curtin is also working on the integration of the distributed authentication infrastructure with their ILMS, Aleph from Ex Libris. Notre Dame University (ND) will be investigating this with their new ILMS in due course. These developments will open the way for other inter-institutional opportunities.

While these works are progressing and the distributed authentication infrastructure is moving to a state of production, other ideas are gestating, waiting for the right time. Perhaps the well established reciprocal borrowing arrangements between four of the universities may benefit from automation of registration. Perhaps the time has come for the negotiation of an extension of the reciprocal rights into electronic usage. Once the distributed authentication infrastructure is established the possibilities that it will enable can be explored and a great step be taken forward in the electronic age.

The Bigger Story

The rich tapestry and the thread that is WALAP

At this point, having described the achievements, difficulties and future hopes for this project it is worth taking a step back and looking at the bigger picture again. The motivations for the WAGUL Authentication Project (WALAP) as they were previously documented (Green 2002) remain current. The issues of authentication, authorisation and access that were enunciated in that paper, and in others published during recent years, remain on the table. A recent submission by the Council of Australian University Librarians (CAUL) to the National Research Infrastructure Taskforce notes that a "further crucial problem for collaborative research work is the matter of access and authorisation management. A solution to the problem of multiple sign-ons to resources across a distributed environment has been difficult to develop" (McPherson 2003, p3). In the years since WAGUL first identified authentication as a strategic issue it is clear that as an industry we are still grappling with issues of Identity Management, authorisation across domains and suitably granular access controls.

Providing a solution to access management relies on a series of interlocking solutions across the domains of identity management, authentication, authorisation and access management (not to mention security, digital rights management, digital asset management, content management, records management and e-commerce). As attention has focussed on each of

these areas the underlying complexity has been revealed. The term middleware has been adopted to describe this complex infrastructure (West 2002) in recognition of the interlocking nature of its components and the role it plays as the middle man in allowing people to get on with their real work. In an ideal world middleware would be largely invisible, always available and never intruding into activities conducted online. However this is a vision that is yet to be reached in this sphere, though aspects of it are within sight.

Identity management can be largely resolved at a university level by the implementation of applications to manage students and human resources. Combined with a whole of university approach and well managed procedures for creation and management of electronic identities these can provide the backbone to identifying the staff and students that comprise the population of a university. Authentication can be addressed by the implementation of enterprise level directory services that provide a 'view' of the identity information and well understood interfaces for authentication, such as LDAP. Authorisation can be addressed at the same time by populating the enterprise level directory services with sufficient information about each identity that authorisation decisions can be made, using the same common interfaces. Access control can then be achieved by applications that have leveraged this infrastructure.

None of this is easy to achieve. Simply implementing a student system can be a major logistical achievement, as has been demonstrated in various universities. Achieving complete, accurate and timely coverage of the student population can be a challenge. Not to mention staff and other persons who have an involvement in the work of a university but whose status can be unclear. If the directory service is dependent on information residing within student and staff systems, as is generally recognised as best practise, then any problems in that area simply flow down the line to other systems. However universities are coming to grips with middleware, as demonstrated at Griffith University (Callow 2003) and finding solutions to the complexity of identity management, as demonstrated at Monash University (Bailey et al. 2003), though such examples simply illustrate the enormity of the task.

The solutions that are starting to be fully implemented tend to be at the institutional level. To be honest there is usually more than enough work in achieving a robust middleware within one's own university without having to think beyond the cloister walls. However the question of domain persists. No university is an island unto itself (or archipelago might be more accurate given the internal complexities of a university). Solutions that provide the ability to manage access within one's own domain may not translate to cross-institutional solutions, and the drivers for cross-institutional solutions are not as strong or immediate as those that are internal. It is clear that collaboration between institutions is part of the political agenda for higher education. This desire for collaboration was clearly enunciated in the recent call for research infrastructure bids by the Department of Education, Science and Training (*Instructions for Institutions Preparing Proposals for Funding under the Research Information Infrastructure Framework for Australian Higher Education* 2003).

The focus on inter-institutional authentication is one of the characteristics of the WAGUL Authentication Project that is innovative. It is focused on inter-institutional authentication at a time when intra-institutional authentication, authorisation and access management are still works in progress in most universities. However it is not unusual that inter-institutional initiatives are being led by the university library sector. Cooperation and collaboration has been a routine part of our business and a focus on authentication is a natural consequence of

the growth in provision of services and resources online, and a desire the share those expensive resources with others.

However the WAGUL Authentication Project is not an isolated instance and there are other initiatives that are interested in the cross-institutional domain.

Shibboleth is a US-based "initiative to develop an open, standards-based solution to the needs for organizations to exchange information about their users in a secure, and privacy-preserving manner" (*Shibboleth Introduction* 2003). In this case the focus is on authorisation and access management rather than on authentication or identity management. Those issues are seen as best managed by an organisation itself, with the identity being 'federated' by use of the Shibboleth software. It should be noted that the focus of the WAGUL Authentication Project is on providing a basis for authentication and authorisation and a tool such as Shibboleth would dovetail nicely with the infrastructure, though this hasn't been demonstrated at this time. In the context of libraries, Shibboleth may become the next generation of the Athens solution which "has been in active use since 1996 in the UK Higher Education community, providing access to many centrally-funded web-based services" (*About Athens n.d.*) and may provide a better long term solution than the widely implemented EZProxy software (*EZProxy Overview* 2003).

The Australian Academic & Research Library Network (AARLIN) is another initiative that is working in the inter-institutional domain that aims to provide a collaborative research information infrastructure that will allow "unmediated, personalised and seamless end user access to the collections and resources of Australian libraries and document delivery services" (*About AARLIN* 2003). While the focus of this substantial project is on access, issues of authentication and authorisation will need to be addressed. It is likely that WALAP will provide possibilities for addressing these issues, though the solution may be multi-faceted given the size of the AARLIN project.

The COLIS project, while demonstrating "collaborative online learning and information services" (*About COLIS* 2003), has also focussed attention on the areas of access management, particularly digital rights management and digital object management. The Digital Objects Repository Management Forum, Sydney in 2003, heard from one of the key players in COLIS (Dalziel 2003) about the need for access management. The interrelated roles that could be played by WALAP (authentication), Shibboleth (authorisation) and Access Management were also mapped out in a way to demonstrate how both intra-institutional and inter-institutional needs could be met and that the ability to work with multiple methods was essential to providing a solution to access management in either domain. The funding of the Meta Access Management (MAMS) Project (McGauran 2003), led by Macquarie University, will provide an exciting opportunity to demonstrate integration of these facets and place Australia at the forefront of this area of development.

In all of these projects, and others, there is a growing recognition that cooperation and collaboration are characteristics of the way forward in the higher education sector. The WAGUL Authentication Project has attempted to provide a solution to one of the persistent barriers to collaboration but the final solution has only begun to be written.

Reference List

About AARLIN 2003, [Online], Australian Academic and Research Library Network, Available from: <http://www.aarlin.edu.au/about.shtml> [10 December 2003]

About Athens [n.d.], [Online], Athens Access Management Services Available from: <http://www.athensams.net/about/> [10 December 2003]

About COLIS 2003, [Online], Collaborative Online Learning and Information Services, Available from: <http://www.colis.mq.edu.au/about.htm> [10 December 2003]

Bailey, N., Jackson, K., Liew, L., Schendzielorz, P., and Treloar, A., & Troeth, L. 2003, 'Implementing a Hybrid LDAP Directory at Monash University to Provide Access to External Users', in *Proceedings EDUCAUSE IN AUSTRALASIA 2003: Expanding the Learning Community - Meeting the Challenges*, Adelaide Convention Centre, Australia, pp. 329-337. Available from: <http://eprint.monash.edu.au/archive/00000023/> [10 December 2003]

Bronleigh, H. 2003, [Online], *Murdoch University Document Delivery*, Available from: <http://wwwlib.murdoch.edu.au/services/docdel/#info> [10 December 2003]

Callow, B. 2003, 'Service Connection through an Enterprise Directory' in *Proceedings EDUCAUSE IN AUSTRALASIA 2003: Expanding the Learning Community - Meeting the Challenges*, Adelaide Convention Centre, pp. 306-317. Available from: <http://www.gu.edu.au/ins/publications/conf_papers/callow03.pdf> [10 December 2003]

Dalziel, J. 2003, 'Access management: challenges and approaches', [Online], *Digital Objects Repository Management Forum*, Sydney, Available from: <http://www.library.usyd.edu.au/dest/dalziel.ppt> [10 December 2003]

EZProxy Overview 2003, [Online], Useful Utilities, Available from: <http://www.usefulutilities.com/support/overview.html> [10 December 2003]

Green, P. 2002, 'Building a shared authentication infrastructure: a matter of trust', In *Proceedings 11th VALA Biennial Conference and Exhibition, 2002: E-volving Information Futures, Victorian Association for Library Automation*, Melbourne Exhibition and Convention Centre, pp. 565-573. Available from: <http://www.vala.org.au/vala2002/2002pdf/40Green.pdf> [10 December 2003]

Green, P. 2002b, [Online], *WA Libraries Authentication Project*, Available from: <http://walap.curtin.edu.au/> [10 December 2003]

Green, P. & Reid, T.A. 2003, 'A Distributed Authentication Infrastructure for Western Australian Universities', In *Proceedings EDUCAUSE IN AUSTRALASIA 2003: Expanding the Learning Community - Meeting the Challenges*, Adelaide Convention Centre, pp. 420-430 Available from: <http://walap.curtin.edu.au/docs/0018anav.pdf> [10 December 2003]

Instructions for Institutions Preparing Proposals for Funding under the Research Information Infrastructure Framework for Australian Higher Education 2003, [Online],

Department of Education, Science and Training, Available from: <http://www.dest.gov.au/highered/research/documents/proposals.pdf> [10 December 2003]

McGauran, P. 2003, '$12 Million For Managing University Information', [Online], Department of Education, Science and Training, Available from: <http://www.dest.gov.au/Ministers/Media/McGauran/2003/10/mcg002221003.asp> [10 December 2003]

McPherson, M. 2003, 'CAUL Submission to National Research Infrastructure Taskforce', [Online], *Council of Australian University Librarians*, Canberra, Australia, Available from: <http://www.dest.gov.au/highered/ri_taskforce/submissions/pdf/r28.pdf> [10 December 2003]

Shibboleth Introduction 2003, [Online], Internet2, Available from: <http://shibboleth.internet2.edu/shib-intro.html> [10 December 2003]

West, A. 2002, 'Middleware: Addressing the Top IT Issues on Campus', *EDUCAUSE Quarterly*, vol. 25, no. 4, pp. 6-10. Available from: <http://www.educause.edu/ir/library/pdf/eqm0241.pdf > [10 December 2003]