

ShibGrid, a Shibboleth based access method for the National Grid Service

David Wallom^{†1}, David Spence^{††}, Kang Tang[†], David Meredith^{††}, Jens Jensen^{††} & Anne Trefethen

[†]Oxford e-Research Centre, University of Oxford

^{††}e-Science Centre, Science & Technology Facilities Council

Abstract

Users access the National Grid Service using individual personal certificates from a trusted Certification Authority. However, the central role of the NGS in UK e-Science makes it desirable that users can use their institutional id to access Grid resources. The UK Federation is deploying a Shibboleth infrastructure to ease access from higher and further education institutions to common resources, and it is thus natural to develop Shibboleth access to the Grid in general, and the NGS in particular. Rather than going through expensive middleware refactoring, we have focused on enabling this via portals and credential conversion.

This project has developed prototype software to allow users to use Shibboleth[2] to access the grid. Long term scalability is improved since authentication is also devolved from the national Grid CAs to users' home institutions. The existing NGS portal was taken and key components developed enabling the capabilities of both Shibboleth and the certificate based authentication to be provided through a user-friendly interface. This required certificate management tools, pluggable security modules for the transfer of Shibboleth attributes, and portlets to manage user's Shibboleth based low assurance certificate.

1. Introduction

Since its inception the NGS has relied upon x509 certificate [3] authenticated access where every potential user needs to obtain a certificate from a certificate authority accredited by the International Grid Trust Federation [4]. Once an account request is reviewed and approved, access is typically granted to all core, and partner sites and the services provided to the NGS by these sites. There is currently no way on the NGS to restrict access to some services but not others, depending on who the user is or which project or organization he is from, other than by managing each individual's authentication separately. This shortcoming can potentially discourage expansion of the NGS to new sites who want to control who can use their resources, especially as the number of users of the NGS increases. This is in addition to the burden of requiring new users to obtain a Grid certificate and to learn how to use it prior to being able to use the NGS.

Shibboleth is a project which allows authorization decisions to be made when a user from an organization requests access to online resources, not necessarily solely as a result of who the person is, but possibly based on other information about the person, such as their role in their organization. Such attributes are maintained by an attribute authority, in this case the user's organization and are only disclosed

with the user's knowledge, so privacy is preserved.

In this way authentication is devolved from the single national entities of the Grid CAs to users' home institutions – this also creates a far more scalable infrastructure where the number of users can increase dramatically. In removing the need for users to obtain certificates for themselves, the process of obtaining access to the NGS is made far easier as certificates and Public Key Infrastructure (PKI)[5] security in general have traditionally been an obstacle for novice users – this problem is becoming more acute as the NGS attracts more users from non computing disciplines.

The advantages of moving to a devolved security infrastructure such as ShibGrid are so significant that the authors believe that it can bring a major contribution to the future success of the NGS. Shibboleth is already proving to be the most popular solution for integrated and uniform access to multiple resources. It is currently being deployed by JISC as a replacement for the very successful Athens service [6].

It is clear that extra attributes applicable to users such as roles within collaborations etc. can be applied using descriptions in SAML etc. They are considered outside the scope of this project which is just looking at a plain Shibboleth – x509 conversion mechanism.

¹ For further information please contact david.wallom@oerc.ox.ac.uk

2. Requirements and Use Cases

The main aim of the project can be summarized as “the removal of barriers inherent in PKI infrastructures which prevent more users taking advantage of the NGS”. This leads to the development of the following four use cases, whose importance to the operation of the NGS is discussed with each case:

1. User with 12 Month Certificate (from an IGTF-accredited Grid CA) creates a proxy certificate and stores it on the ShibGrid MyProxy server (accessed via the NGS portal) using only his Home Organisation’s (HO’s) single sign-on (SSO) for authentication. The user can thereafter access and use the proxy certificate using only his HO’s single sign-on for authentication.

This use case is important for existing users of the NGS and necessary for a successful deployment and initial adoption of ShibGrid.

2. User who does not own a 12 Month Certificate is able to perform some Grid functions (for which at least a lower level of assurance certificate, or its associated proxy certificate, is necessary) via the NGS portal and using her HO’s SSO, but without needing to apply for a 12 Month Certificate.

For the large number of potential users of the NGS who do not need the additional assurance that a 12 Month Certificate provides, this use case enables the number of NGS users to scale significantly into the future.

3. Users must be able to be registered and approved as NGS Users and to be removed/revoked from the list of NGS Users.

It is essential that remote resource owners maintain the ownership and control that they currently have over the resources they have. This will reassure possible NGS partners and help maintain the increase in numbers of institutional partners.

4. A user that has generated a Temporary Certificate should be able (if he so wishes) to take control of that certificate and use it in a similar manner to a 12 Month Certificate.

Within the NGS the certificate is used for other areas outside pure job submission etc. These include NGS website access using GridSite[7] etc.

In addition to these use cases it was clear that there would be additional requirements that we would have to satisfy.

The ShibGrid project sits between the two worlds of Shibboleth and Grid/NGS. It must therefore satisfy the standards of both, primarily Shibboleth/SAML and the Grid Security Infrastructure (GSI) [8]. It must also additionally track changes to these standards and technologies that are employed. In the case of Shibboleth this is the change to SAML 2.0 and Shibboleth 2.x, along with the formulation of the UK Federation policy [9]; and in the NGS this is the adoption of VOMS [10]. The project must also provide input into the formulation of the UK Federation policy to attempt to ensure that it stays compatible with ShibGrid. Since the Shibboleth-Enabled Bridge to Access the NGS (SHEBANGS) [11] project has an authorization focus and is looking in detail at integration with VOMS, the ShibGrid project did not look at VOMS in-depth.

Security was also seen as a key concern in ShibGrid and so it was decided to add tests to ensure the security of the MyProxy server.

3. Implementation

DIAMOND[12] and Integrative Biology[13] were chosen as primary users, partly because they are represented at the University of Oxford and STFC, so we could work closely with them if necessary, partly because they were not using the Grid already.

3.1 User and Stakeholder Requirements

DIAMOND already had well defined use cases for single sign-on, and we used these for ShibGrid. In the case of IB we interviewed their users and translated their feedback into requirements that are relevant to ShibGrid.

Other stakeholder requirements (for the NGS, UK e-Science CA[14] and OMII[15]) were derived from the relevant public documents from, and discussions with, the organization.

The results of the User Requirements exercise can be found in the document at <https://wiki.oerc.ox.ac.uk/shibgrid/Documentation>. They are presented here to provide the background for the description of the architecture of ShibGrid.

In summary, the requirements gathered from OMII provided our methodology for code and documentation; the requirements from the NGS and the UK e-Science CA largely defined the rules within which the ShibGrid architecture must operate; and the requirements from

The steps are as follows:

1. User requests access to the NGS portal, a Shibboleth Service Provider (SP), through a Shibboleth logon, and the user's browser is then redirected to the Where Are You From (WAYF) service.
2. The user chooses the appropriate Identity Provider (IdP) from the form returned by the WAYF.
3. The user's browser is re-directed to the correct IdP.
4. The user is authenticated by the IdP SSO service through the institutions authentication mechanism.
5. The IdP redirects the user's browser back to the portal (the Service Provider or SP in Shibboleth terminology). The signed authentication SAML assertions are passed in this redirect.

The portal calls out to the IdP's attribute authority for attributes about the user.

6. The username is extracted from the (Base 64-encoded) signed attribute assertion. The extracted username and (Base 64-encoded) signed attribute assertions (used as the password) are sent to the MyProxy server. The ShibGrid MyProxy server returns either a proxy of the user's certificate (if the user has an e-Science certificate and has already uploaded a proxy) or returns an automatically-generated low-assurance certificate from MyProxy's built in CA, if they have not uploaded a proxy.

7. The certificate or proxy certificate is returned to the portal.

8. The user can access the NGS via the portal.

Steps 1-5 are the same as standard Shibboleth access, except that the SP requires a signed attribute assertion rather than an unsigned one. This is standard option in the Shibboleth federation metadata. Therefore the Shibboleth federation and the user's IdP do not need to be modified.

The use of MyProxy in steps 6 & 7 does not modify the standard MyProxy protocol and is very similar to what a standard portal would do. Therefore this means that portal integration should be reasonably straightforward and comprehensible to portal developers. This also simplifies the method used to access a portal for a first time user (without a certificate), so it becomes the same as other Shibboleth SPs, there is no need to pre-register anywhere before accessing the portal.

A similar method is used to download and upload proxies to and from the ShibGrid MyProxy server. In these cases the functionality is split between code running on

the web server and on the client's machine (an applet in the user's browser) as the system requires access to resources (like files and certificates) on the user's machine.

4. Project Outputs

The project's main outputs were in the form of code and its associated documentation. Here we describe each code component in turn.

4.1. MyProxy server changes

The central component in the ShibGrid system is the MyProxy server. The modifications to the MyProxy server consisted of a new authentication method and a new type of username-to-DN mapping. Both these components are semi-modularised within the MyProxy source. The patch consisted of numerous minor changes to the source to hook in the new methods and pass through configuration options which can be en/disabled at compile time or runtime and are fully configurable, including the format used to convert Shibboleth attributes to DNs.

Development of this component also involved implementing parts of the Shibboleth library APIs that are not implemented in the actual SP distribution because they are not used.

We considered three choices for DN mapping schemes for low-assurance certificates in ShibGrid. They were as follows:

1. `/C=UK /O=eScienceMyProxy /OU=<Institution>/UID=<Site username>/CN=<First name> <Last name>` This scheme has the advantage that it is very close to the DN format currently in use for the UK e-Science CA and gives resource administrators a good indication of whose certificate it is. The disadvantage is that within the UK Shibboleth Federation none of the attributes required for this DN format are released as standard by IdPs and with outsourced IdPs many do not have the access to these attributes.
2. `/C=UK /O=eScienceMyProxy /L=<IdP entity-id>/CN=<eduPersonTargetedId>` This scheme has the main advantage that `eduPersonTargetedId` is a core UK Shibboleth Federation attribute so there would not have to be any negotiation with sites over data protection issues. Concerns were raised over the traceability of users through this form of DN as the IdP logs will be the only place which will provide a mapping back to the user (these logs might be on the server of a commercial IdP provider). Crucially, because the user can

obtain Grid credentials when logging on through many different SPs (e.g. portals, download tools) this scheme cannot be used as each SP will be sent a different eduPersonTargetedId.

3. `/C=UK` `/O=eScienceMyProxy`
`/CN=<eduPersonPrincipleName>` This scheme has the advantage that it provides some clues to the resource administrators who is behind the certificate and it is feasible in most cases that traceability would be possible without access to the IdP logs. The eduPersonPrincipleName attribute is a recognised UK Shibboleth Federation attribute, although not a core attribute. This means that some liaison with home institutions over data protection issues may be required. All SPs receive the same eduPersonPrincipalName which means that the problems of eduPersonTargetedId do not affect this scheme. A minor technical complication is the encoding because the ePPN may contain an '@' sign which cannot be encoded as a printableString. Different alternative encodings are possible but not all work equally well with middleware.

While early testing, both within our own federation and within the SDSS federation, we used Scheme 1. When we later wanted to involve other institutions to use ShibGrid for testing purposes, we had to change over to Scheme 3 which is now our preferred scheme. If a later policy change by the UK federation or other key parties means a different DN scheme becomes the best solution then the ShibGrid system can easily be reconfigured to use a new format as this is all set in configuration files.

It should be recognised that we cannot use the same scheme between our Shibboleth based low assurance CA and the main UK e-Science CA as it would be very easy for both the system itself, users and resource owners to become confused by making differentiation between low and medium assurance certificates difficult.

4.2. Registration Page

The registration page performs three functions:

1. Checks compatibility with a user's IdP.

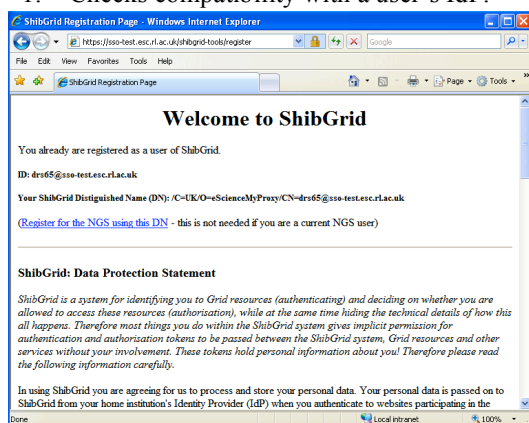


Figure 2 The ShibGrid registration page

2. Provides a platform for user to agree to the ShibGrid data protection policy (which is required by the UK Shibboleth Federation due to the way we use attributes) and any other policies (e.g. acceptable use policy).
3. Provides a link to the NGS registration system with the users DN already entered (and a way for advance users to discover their ShibGrid DN).

Once the user has successfully logged in via Shibboleth to the registration page they are asked to agree to the ShibGrid policy. When they agree to the policy they are shown their DN (but they do not need to remember it) and a link to the NGS registration page so they can register without needing to handle their DN. A screenshot is shown in figure 2.

4.3. Hardened ShibGrid-enabled Stringbeans Portal

The primary portal to be ShibGrid-enabled in this project was Stringbeans portal framework 2.4.1, on which the older NGS portal was based. ShibGrid-enabling this portal built on work done at CCLRC-Daresbury on the NGS portal, in particular work to enable login from a MyProxy server, and work done at Oxford in the SPIE[18] project in Shibboleth-enabling

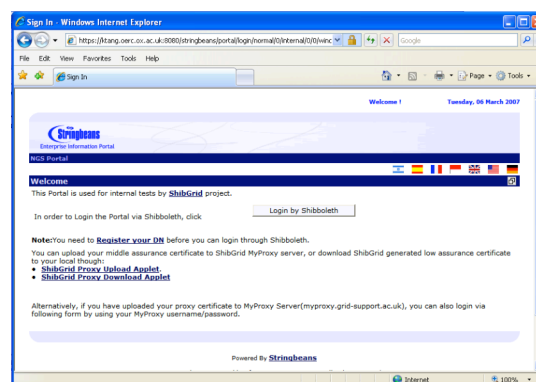


Figure 3 The Shibboleth enabled Stringbeans portal

portal frameworks.

The SPIE project has developed a JAAS (Java Authentication and Authorisation Standard) module to enable Shibboleth login to general portal frameworks. This module reads in users SAML attributes through the SP component that sits in front of the portal framework. The challenge in the portal work is to obtain the full attributes assertion from the user and then pass it through to a MyProxy JAAS login module, which in turn uses this assertion as password to authenticate to a remote ShibGrid Myproxy Server. Unlike the

login process in SPIE, a ShibGrid login is only successful if both steps succeed.

Another important issue which was solved was how to provide users with the ability to renew expired Grid credentials within the portal, as this may require a user to perform another Shibboleth login. So, as with most of the components a simple interface (i.e. a “Login via Shibboleth” button) hides a lot of technology.

4.4. Prototype ShibGrid-enabled uPortal portal

Towards the end of the project it became possible to ShibGrid-enable the newer uPortal portal framework on which the newer NGS portal is based. The key issue here is that unlike Stringbeans, uPortal does not support the JAAS module standard for authentication, which is used in the majority of portal frameworks. We adopted a solution from the SPIE project, to implement a Shibboleth SecurityContext to invoke SPIE's Shibboleth JAAS LoginModules. Other than this, the implementation is very similar to that used in Stringbeans.

The most important result of achieving uPortal integration is not in tracking developments of the NGS portal, but in providing another ShibGrid-enabled portal framework and another example for current and future Grid portal developers of how to ShibGrid-enable their grid portals.

4.5. Proxy certificate download tool

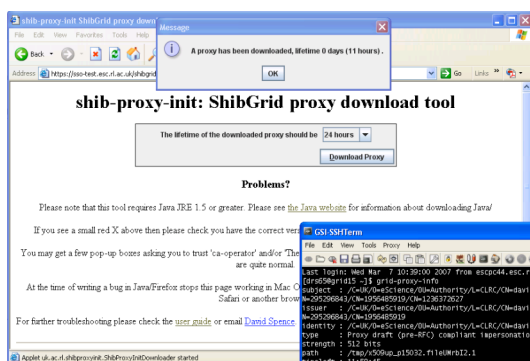


Figure 4. The ShibGrid proxy download tool

The proxy download tool needs to have both a presence on the web server (so it can authenticate users using Shibboleth) and on the user's machine (so it can write the downloaded proxy to disk). Therefore the download tools consists of two Java components: an applet that is displayed within the user's browser and a servlet running in a Tomcat instance, protected through a Apache Shibboleth SP.

The user interface is very simple with a drop-down list to select the desired lifetime of the proxy and a “Download” button. When the user clicks the “Download” button, a proxy request is sent to the servlet to which the user has already logged onto via Shibboleth. The servlet can then go ahead and request a proxy certificate from the MyProxy server, returning it to the applet. The private key for this proxy certificate is never transmitted across the network. The applet then writes the proxy certificate to disk. This is shown in figure 4

4.6. Proxy certificate upload tool

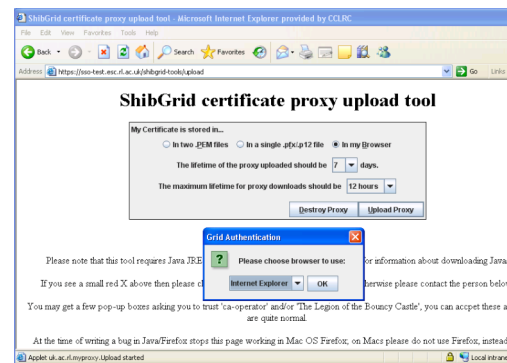


Figure 5. The ShibGrid proxy upload tool

The same split between applet and servlet is seen in the proxy certificate upload tool. This time the interface is slightly more complex, but it boils down to specifying where your certificate is to be found (in your browser, in a back-up file, or in a normal Grid setup) and specifying a policy for the uploaded proxy certificate (how long it should persist and the maximum lifetime of proxy certificate that the MyProxy server will generate from it). Users can use this tool to upload proxy certificates or destroy already uploaded proxy certificates

When the user clicks the “Upload” button, then the applet requests the user's attributes from the servlet, which also delegates its right to use these attributes to the applet. The applet then can upload the proxy certificate to the MyProxy server. A screenshot of this is shown in figure 5.

5. Outcomes

The ShibGrid project has created a Grid authentication solution, using Shibboleth technology, which satisfies the needs of those users, and it enables the use of a variety of Grid access methods (portal, GSI-SSHTerm, local Grid clients, etc.,) out of the box.

The project had two criteria for delivery: production level code and associated

documentation fulfilling the project use cases and the objectives.

There were many areas where we tried to ensure that the project's code was of production quality, the adoption (where possible) of OMII standards, peer review of code, management review of documentation, automated and human testing and user testing. The result is a robust and secure system which the project management team is confident is production-ready. Automated tests have been developed to check the implementation's conformance to the security aspects of the architecture.

The project use cases were fulfilled as follows:

Use case	Components
PUC1	MyProxy server, Upload tool, Download tool, Portal
PUC2	MyProxy server, Portal
PUC3	User Registration Page
PUC4	MyProxy server, Download tool

Going back to the Objectives from section 5 (Implementation, under User Requirements):

No	Descr.	Status
1	Allow users with no prior knowledge of PKI to use NGS	Done, by ShibGrid enabling the NGS portal
2	Allow users with e-Science certificates to use the system	Done via MyProxy and associated upload tool
3	Provide stable mapping in DN presented to resources	Done by implementing appropriate naming scheme
4	Users who change institution should have same access	Done by implementing appropriate naming scheme
5	Resource admin should see user information	Met as best we could under constraints imposed by UK Fed. policy

6	Users apply for access using normal procedures	Done – users are forwarded to the normal registration page with their ShibGrid DN
7	Work with NGS operations for users with more than one DN	Handed over to NGS ops. Workaround available.

This latter objective (number 7) is not fully met at the time of writing, but we expect it will become higher priority for NGS operations once ShibGrid-enabled services are wider deployed. The consequences of not meeting this objective are slight, though: few users will ever have more than one DN: we ensure consistent DN allocation by having chosen naming scheme 3, using the eduPersonPrincipalName. Those that do have more than one DN can be managed (renamed and remapped) manually by NGS support.

When ShibGrid is fully deployed, along with a fully deployed UK Federation, it will have a major impact on the take-up of Grid computing by those with no desire to learn the peculiarities of the Grid Security Infrastructure, now there is an easy to use way to get basic and advanced Grid access through a technology which all academics and students will soon become familiar with. This will mean better, faster research results as researchers either using Grid resources where they would not have before or have to spend significantly less time learning more technology freeing up time for core research.

Conversely, ShibGrid will contribute in a large way to attracting large numbers of users from all areas of research to the NGS, as long as their work can be done via the portals, or their applications are otherwise Grid-enabled.

Deployment of ShibGrid could also give some impetus to the adoption of Shibboleth as it provides a compelling application (i.e. free at point of use computing resources) through Shibboleth that was not available in the previous Athens authentication framework.

Along with the SHEBANGS project, the ShibGrid project has also had an impact in ensuring that the UK's interests are represented in the forums discussing standards for Grid and Shibboleth integration and that the needs of Grid systems were represented in the formation of the policy of the new UK Shibboleth federation.

It is always a risk for a project to rely on external projects, and we found the risks

materialised with the external user testing. DIAMOND and IB were, for all sorts of good reasons, not able to provide the time for testing, especially since they do not have many, if any, end users yet who can actually do this. Even so, it was still invaluable to engage these (longer-term) projects in the area of user requirements, as obtaining meaningful user requirements independently in a one year project would have been a difficult challenge.

6. Conclusion

The ShibGrid project was a success using existing components, Shibboleth, MyProxy, and standard portals, producing components to combine the UK Shibboleth and NGS infrastructures to enable users, from novice to expert, to access the NGS. It meets the requirements of initial users, DIAMOND and Integrative Biology, for single sign-on and Grid access. The prototype was deployed in the NGS portal but not tied to a single portal - we managed to ShibGrid-enable both uPortal and StringBeans-based portals. Via credential upload and download tools we have enabled portals as well as "thick" clients (desktop applications, normally outside the reach of Shibboleth) and "thin" clients (e.g. browsers) to benefit from the credential management built in this project.

Security issues that were addressed include passing signed assertions between components, addressing identity and anonymity/pseudonymity issues, as well as resource logging and access auditability.

The impact of this project spreads over many areas: first and foremost, it succeeded in meeting its primary goal, to provide Shibboleth access to the Grid. This will enable the NGS to grant access to users who do not have, and do not want, an e-Science certificate, thus lowering the barrier for beginners, and widening the user base. Furthermore, using standard components and protocols ensures the product is easily deployable, maintainable, and interoperable.

7. References

1. UK National Grid Service, Details available at <http://www.grid-support.ac.uk/>
2. T.Scavo and S.Cantor. Shibboleth architecture technical overview. Internet 2 documents:draft-mace-shibboleth-tech-view-02, June 2005.
3. RFC 2459, The x509 certificate, IETF, <http://www.ietf.org/rfc/rfc2459.txt>, January 1999.
4. International Grid Trust Federation (IGTF), Details available at <http://www.gridpma.org>
5. Public Key Infrastructure (PKI), An example definition located at http://en.wikipedia.org/wiki/Public_key_infrastructure
6. Connecting People to Resource, JISC, http://www.jisc.ac.uk/uploaded_documents/JISC-BP-ShibFedAcc-Inst.pdf, February 2006
7. Gridsite, gacl and slashgrid: Giving grid security to web and file applications, AT Doyle, SL Lloyd, A McNab - Proceedings of UK e-Science All Hands Conference, 2002
8. Security for Grid Services, V Welch et al, Proceedings of the 12th IEEE Symposium on High Performance Distributed Computing, June 2003.
9. UK Shib Federation
10. From gridmap-file to VOMS: managing authorization in a Grid environment, R Alfieri et al, Future Generation Computer Systems, 2005.
11. SHEBANGS (Shibboleth Enabled Bridge to Access the National Grid Service). <http://www.sve.man.ac.uk/Research/AtoZ/SHEBANGS>, June 2006
12. DIAMOND, Details at <http://www.diamond.ac.uk>.
13. Integrative Biology - exploiting e-Science to combat fatal diseases. D. Gavagan et al, e-Science All Hands Meeting, September 2004
14. UK e-Science Certificate Authority, Details at <http://www.grid-support.ac.uk/content/view/23/27/>.
15. Open Middleware Infrastructure Institute UK, Details at <http://www.omii.ac.uk>.
16. StringBeans by NABH, <http://www.nabh.com/projects/sbportal>
17. uPortal by JA-SIG, <http://www.uportal.org/>
18. SPIE (Shibboleth-aware portals and information environments) Details at: <http://www.oucs.ox.ac.uk/rts/spie/>