PROCEEDINGS OF SCIENCE

Linking Authenticating and Authorising Infrastructures in the UK NGI (SARoNGS)

Mike A.S. JONES[*†]
The University of Manchester
E-mail: mike.jones@manchester.ac.uk

Jens JENSEN
Science and Technologies Facility Council
E-mail: jens.jensen@stfc.ac.uk

The UK NGS aims to provide simple trusted access to digital services for the UK's research community, in particular but not limited to Grid and Cloud provision. To achieve this we have to satisfy conditions laid down by three types of entities: individuals, resources, and the identification and attribute authorities who vouch for them. We have to set the bar high enough to satisfy resource owners, low enough to let most legitimate users in and yet also satisfy legal requirements. This makes it difficult if not impossible to fit one access mechanism to all stakeholders.
SARoNGS was a JISC funded technical project that was developed in the UK to apply a federated access model (The UK Access Management Federation for Education and Research, based upon Shibboleth) to the grid environment. It resulted in a production service supported by the UK NGI to issue grid credentials, obtain Virtual Organisation Membership Service (VOMS) assertions and place them within reach of the user so to provide access these online digital services.
We present the details of this service, the ways we joined the loose ends together, the remaining issues and future directions.

EGI Community Forum 2012 / EMI Second Technical Conference,
26-30 March, 2012
Munich, Germany

[*]Speaker.

[†]Paper presented on behalf of the UK NGS. We would also like to acknowledge A.J. Richards *et.al.,* members of the initial SARoNGS project [1].

PoS(EGICF12-EMITC2)150

1. Introduction

The SARoNGS service currently sits between two infrastructures that make use of separate authentication authorisation mechanisms: The UK's National Grid Service (NGS) which relies upon the Grid Security Infrastructure [2]: a Public Key Infrastructure modified to simulate requirements of grid computing[1], and the UK Access Management Federation for Education and Research [3] (UK AMF) which is a Shibboleth based infrastructure primarily designed to provide access to library resources (and similar) to members of UK academic institutes. The SARoNGS service issues pseudo-anonymous[2] X.509 grid credentials to members of institutes which are subscribed to the UK AMF. These credentials are made available through MyProxy in the form of GSI proxy certificates and may optionally contain VOMS assertions. The private key material associated with the initial X.509 certificate itself is kept only for a short period and never released by the service.

SARoNGS was designed to allow individuals from all research community backgrounds to access grid resources without needing to understand the complexities peculiar to the identity management technologies chosen by the grid development community. It made the following assumptions:

- Individuals accessing resources would be registered at institutes subscribed to the UK AMF;
- Individuals use browsers;
- The research communities provide grid portals.

In essence, after being referred to and at the bidding of an individual, SARoNGS will create an authentic, traceable, grid credential (including VOMS assertions) representing that individual. To establish trust during this process, the individual must log into a recognised institute and that institute must releases a signed digital statement describing that login including some basic authentication information to the SARoNGS service. The resultant grid credential will be made available to the web-portal that the individual used to initiate the login process[3].

Conventional portals usually maintain their own register of individuals. These users will often have to log in by passing username/password challenge. Those portals then perform activities on the users behalf using, where necessary, identity credentials belonging to the portal not the user. In this scenario, the portal has a number of policy driven limitations restricting what it can do especially within a grid environment [4]. SARoNGS addresses these policies and reduces the limitations by providing strong delegatable identity assertions which the portal can use. Further, it negates the necessity to maintain a portal specific user registration system.

Due to a number of other complex policies e.g.: data protection law and the users' privacy requirements, policies laid down by and to protect the UK AMF, registration requirements of the Virtual Organisations[4] (VO), the operational requirements of the resource providers (and by impli-

[1] In particular, the NGS relies upon Single Sign-On (SSO), and Delegation (identity delegation, or impersonation in this case).

[2] The pseudo-anonymous credentials are issued not by design of the SARoNGS project but by policy determined by the infrastructures SARoNGS relies upon.

[3] This is known as the *portal first approach*. The individual is also able to initiate the whole login process directly and outside the scope of a web portal, in which case that individual will be presented with instructions on how to obtain a delegated credential for themselves using grid middleware, and may present these details manually to a web portal of their choice.

[4] Virtual Organisations essentially are groups which map onto user communities. See e.g. [5].

cation policies governing trust infrastructures defined to support them, e.g. the *International Grid Trust Federation [6] (IGTF)*), the SARoNGS approach has had to overcome a number of technical and political obstacles. We present these obstacles and, where we have found it possible to solve them, those solutions.

1.1 History

In 2006 a Bird-of-a-Feather session was scheduled at the then *Global Grid Forum* with the purpose of discussing grid and Shibboleth interoperation [7]. At this meeting a number of projects were presented. The outputs of two of these: ShibGrid [8] and SHEBANGS [9], subsequently merged to form the SARoNGS project and ultimately the SARoNGS service for the UK NGS.

1.1.1 SHEBANGS

SHEBANGS focused on the release of GSI proxy certificates[5] based upon a SAML assertion received through an authentic Shibboleth Login process. The service would create a short-lived[6] End Entity Certificate (EEC) on the fly using a local online CA and derive a GSI proxy certificate subsequently to be placed in a MyProxy [11] service elsewhere. The demonstrator also developed functionality to create and insert attribute assertions and LoA[7] statements into these GSI proxies before release. The form of the attribute assertions was identical to those provided by a VOMS[8] service [14]. The resulting SHEBANGS Credential Translation Service (whose basic *CTS first*[9] schema is shown in figure 1) assumed several roles:

- A *Shibboleth Service Provider* to manufacture grid credentials for authenticated users,
- A *Certificate Authority* using SAML assertions in place of the usual CA identity vetting process,
- An *Attribute Authority* using both attributes derived from the Shibboleth context and attributes known by the service about the user.

It was originally envisaged that an instance of the SHEBANGS CTS would be run for or by each VO. Portals would, like Shibboleth *Service Providers*, offer login via VO instead of an IdP directly. Like the Shibboleth WAYF service there would be a similar discovery mechanism for the users' VOs: (WVOAYF), see figure 2.

[5] These are credentials derived from an X.509 certificates used in grid computing to provide single sign-on and identity delegation (the GSI proxy certificate specification was later formalized into an IETF RFC [10]).

[6] Short-lived is generally accepted to be about a week or 1,000,000 seconds (approx. 12 days)

[7] Level of Assurance statements based on those defined by NIST [12] were actually conveyed, not through the EEC itself, but through the choice of Certificate Authority used to sign the EEC. The CA's CP/CPS reflected the relevant NIST LoA values. To consume these the relying parties just needed to install those CAs or use grid *signing_policy* files to select the LoA they required for authentication to their systems.

[8] The Virtual Organisation Membership Service (VOMS) had been developed by the European DataGrid [13] as a means of moving from a synchronized, cached, list based authorisation scheme (where all service were required to maintain a current list of authorised users' X.500 Names by polling an authoritative service) to a push based assertion approach.

[9] There are two methods to invoke SARoNGS authentication, *CTS first* and *Portal first*, the later is invoked when a portal refers the user to the CTS with a return URL.

1.1.2 ShibGrid

ShibGrid focused on the release of strong X.509 credentials using a SAML [16] assertion[10] also obtained via Shibboleth and subsequently provided to an online Certificate Authority via the MyProxy Protocol [18]. The strength behind these X.509 credentials was based upon its architecture and the checks and cross-checks performed by the software:

- The public facing service (a Shibboleth Service Provider) and the online Certificate Authority (CA) key signing service were separate network identities, the latter protected by a strong firewall.
- The key material for the online Certificate Authority's key signing service was generated and stored within a Hardware Signing Module[11].
- Connection to the online CA was via the MyProxy protocol, over Transport Layer Security (TLS) and mutually authenticated using a named IGTF based X.509 certificate belonging to the public facing service.
- The SAML assertions exchanged between the public facing service and the user's institutional Identity Provider were passed to the online CA over the MyProxy protocol.
- The authenticity of the SAML: assertions were required to be signed by the IdP and targeted at the public facing service. The *Issuer, Audience, Attributes* and *Signature,* were extracted and verified.
- Authenticity of the Issuer was derived by signature verification on the SAML assertions with the public key material in the UK federation metadata[12].
- A cross check between the network identity of the incoming request and the Audience of the SAML assertion was made by comparing key material used in the former with that found in the Audience's metadata key material entry.

1.1.3 SARoNGS

The SARoNGS project [1] merged ShibGrid and SHEBANGS and established a pair of demonstrating consumer applications: The NGS Portal [19] was equipped with a *Login via the UK Federation* Mechanism, MIMAS produced a Federated login mechanism to a geo-spatial service Landmap [20] with aspiration to employ grid services for data processing. As ShibGrid and SHEBANGS were merged the original outputs were refined by:

- Substituting the weaker CAs of SHEBANGS with the Stronger CA from ShibGrid,
- Removing the LoA assertions[13],
- Removing the dependency on PERMIS, to performing simpler attribute translations locally,
- Replacing the WVOAYF with a VOMS query mechanism allowing multiple VOMS assertions to be obtained placed in the subsequent GSI proxy certificate.

[10] It required a signed SAML Attribute Statement.

[11] HSMs are designed to allow keys to be stored where they cannot be subsequently retrieved, see e.g. [17]

[12] This is currently published regularly at http://metadata.ukfederation.org.uk/ukfederation-metadata.xml.

[13] LoA was not part of the ShibGrid CA and LoA had also largely been ignored by the grid community.

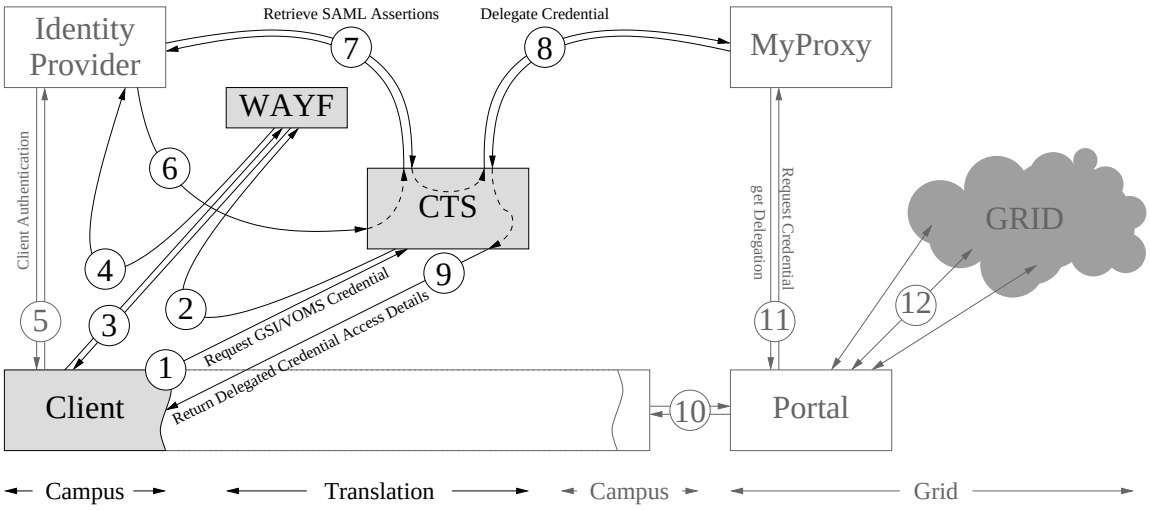

Figure 1: CTS First SHEBANGS

The client accesses the CTS to obtain a credential. Afterwards they can access any suitably configured portal using the resulting MyProxy credential access parameters.

1. The client points their browser at the CTS web server.
2. The CTS requires a login; it uses Shibboleth to redirect the client's browser to a "Where are you from" (WAYF) service.
3. The client selects their home institute from a list, the WAYF. This is usually maintained by the federation in this case the UK AMF.
4. The WAYF redirects the client's browser to their home institution's login facility.
5. The client authenticates with their home institute.
6. The Home Institute redirects the client's browser back to the CTS with either a signed SAML assertion: *Browser POST Profile* or a means to obtain one: *Artifact Profile*.
7. If required (e.g. when using the Artifact Profile), assertions can be retrieved over a back channel.
8. The resulting SAML is processed and a PERMIS [15] service is called to make a policy based decision on whether to release a credential for the client. The CTS then delegates a grid credential over a back channel to a MyProxy server.
9. The CTS presents MyProxy login details, with which to access the grid credential, back to client.
10. The client logs into portal manually providing the MyProxy login details.
11. The Grid Portal supplied with MyProxy credentials uses these to obtain a delegated grid credential. It may then use the grid credential to authenticate the client.
12. The Grid Portal accesses the grid on behalf of the user using the delegated credential.

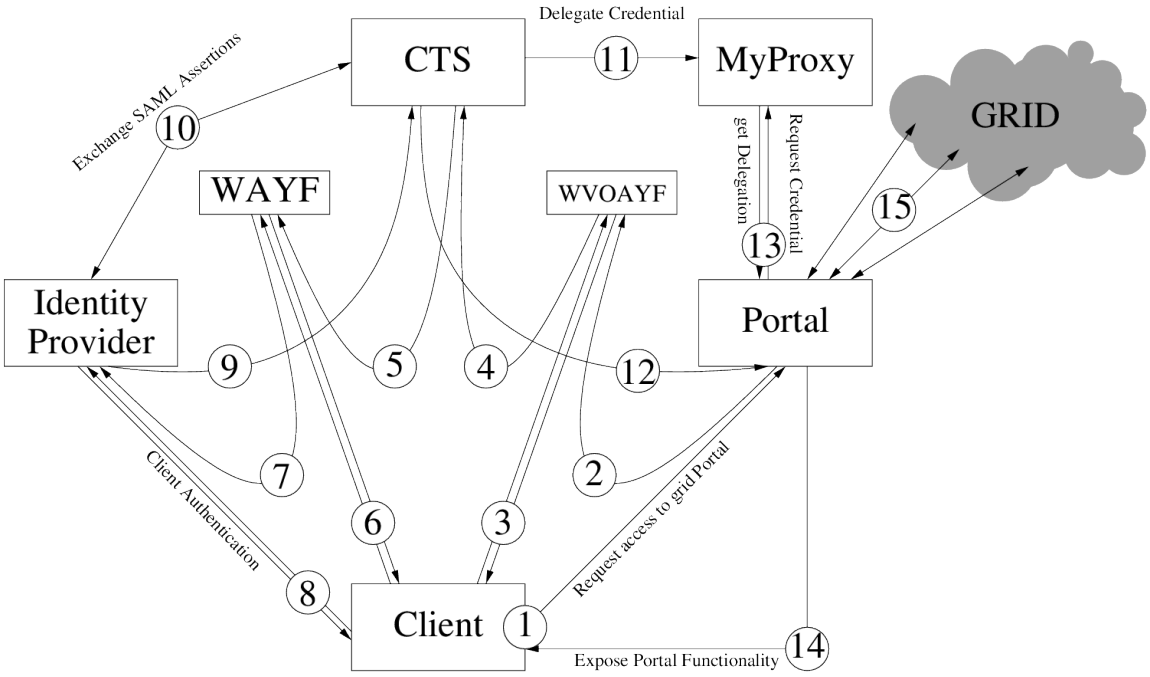

Figure 2: Portal First SHEBANGS

The client accesses a portal and is directed to login via a CTS, the resulting MyProxy credential access parameters are passed back to the portal.

1. A request is made to login to the portal.
2. [optional] The client's browser is redirected to the WVOAYF service[a] (if there is more than one CTS).
3. [optional] The client select their VO (or CTS) from a list.
4. The Client's browser is redirected to the CTS.
5. Immediately the browser is redirect to Shibboleth WAYF service.
6. The client selects their home institution from a list.
7. The WAYF redirects the client's browser to their home institute's login facility.
8. The client authenticates with their home institution.
9. The home institution redirects the client's browser back to the CTS with either a signed SAML assertion: *Browser POST Profile* or a means to obtain one: *Artifact Profile*.
10. If required (e.g. when using the Artifact Profile), retrieval of assertions over a back channel.
11. The resulting SAML is processed and a PERMIS [15] service is called to make a policy based decision on whether to release a credential for the client. The CTS then delegates a grid credential over a back channel to a MyProxy server.
12. The CTS redirects the clients browser back to the portal. The redirection provides the portal with MyProxy login credentials with which to access the grid credential.
13. The Grid Portal supplied with MyProxy credentials uses these to obtain a delegated grid credential. It may then use the grid credential to authenticate the client.
14. The Grid Portal exposes grid functionality to the user via the web interface.
15. The Grid Portal accesses the grid on behalf of the user using the delegated credential.

[a] This stands for Which VO Are You From.

The VOMS query mechanism required the development of an additional registration mechanism to interface with the VOMS service, due to the end entity certificate's key material no longer being directly in the possession of the user[14], see section 4.4 for more details.

2. Implementation

As described above the SARoNGS service is comprised of two services a "Shibboleth enabled" MyProxy service[15] and a Shibboleth Service Provider: the Credential Translation Service (CTS). After the CTS obtains the End Entity Certificate (EEC) from the "Shibboleth enabled" MyProxy service it generates and delegates a GSI proxy certificate from it to the to the UK NGS's MyProxy service[16] for the user (or portal, acting on behalf of the user) to access[17]. This current architecture (with simplified steps) is shown in figure 3.

2.1 Shibboleth Enabled MyProxy CA

The MyProxy CA operates as described in Section 1.1.2. There are two operational differences between this service and the a normal MyProxy Service. Firstly, it operates in CA Mode, issuing short-lived X.509 End Entity Certificates to authorised network entities rather than storing and delegating GSI proxy certificates for the general community. Secondly, it consumes SAML assertions during the issuing process: a modification designed specifically for ShibGrid and SARoNGS.

The SAML assertions are expected to carry the attribute eduPersonTargetID (ePTID) [21]. It is this attribute that will form part of the the Distinguished Name (DN) of the issued X.509 certificate[18] after disguising it via a hashing algorithm.

The MyProxyCA issues EECs within the following CA hierarchy[19]:

1. `/C=UK/O=eScienceSLCSHierarchy/OU=Authority/CN=SLCS Top Level CA`
2. `/C=UK/O=eScienceSLCSHierarchy/OU=SARoNGS/CN=NGS Shib SLCS`

The first is the DN for a self-signed certificate which was used to sign the MyProxy CA certificate, belonging to an offline CA, and which will revoke the MyProxy CA if the service is compromised. The second is the DN of the MyProxy CA itself, which in turn signs the certificate requests from the

[14] This is an assumption made by the VOMS Admin service vendors.

[15] This is a MyProxy server modified to consume and verify SAML assertions.

[16] This second MyProxy service is outside of the scope of SARoNGS, it is an integral component of the NGS infrastructure.

[17] At no point does the CTS place a grid credential directly into the user's browser.

[18] The use of ePTID is not ideal as it contributes to difficulties in the CA's IGTF accreditation, see section 4.7. Furthermore, the attribute value depends not only upon the individual but also the Service Provider. Therefore, having multiple CTS instances controlled by different organisations will produce different credentials and this in turn will cause the CA to issue differently named certificates. This could be solved bilaterally for each IdP-SP relation by requesting ePPN to be released and used instead, however, this would negate the benefits of using the established UK Access Management Federation for Education and Research.

[19] Until recently this hierarchy was part of the UK eScience Certificate Authority hierarchy, but due to deficiencies in the installation tools used by one grid community it was not possible to install the SLCS Top Level CA along side IGTF CA distributions. After separation: making the SLCS Top Level CA a *Root* or *Self Signed* Certificate this became possible.

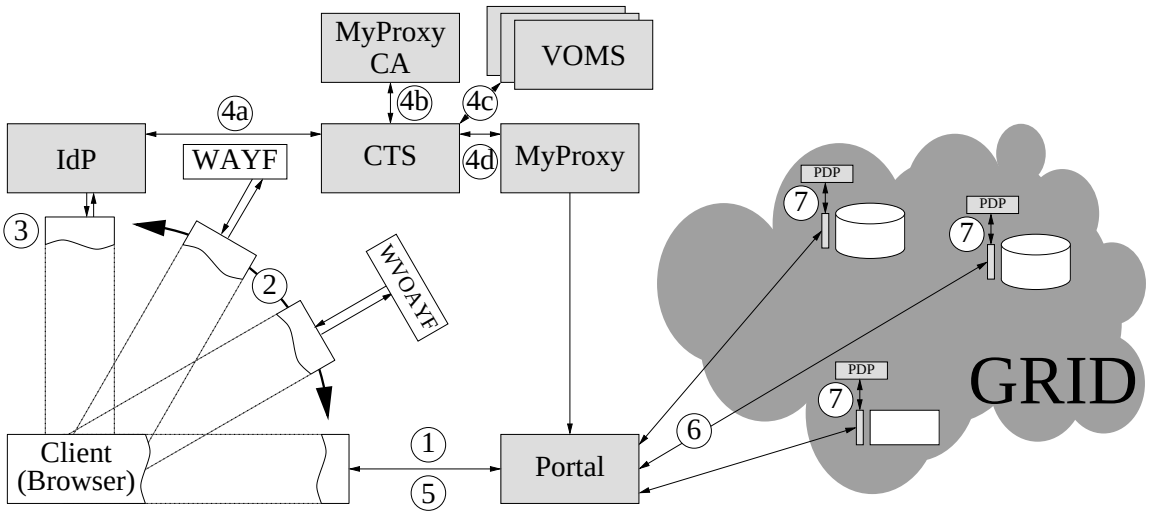

Figure 3: SARoNGS Architecture

(The CTS VOMS registration mechanisms are not shown here.)

1. As before (figure 2), the client accesses the portal first.
2. The client is redirected to their home institution's login page via the CTS and WAYF services, as before.
3. The client logs in to their home institution.
4. a, The client is authenticated to CTS as before.
   b, The CTS requests a certificate from the online CA (MyProxyCA).
   c, VOMS credentials are obtained using the VOMS GSI protocol
   d, The GSI proxy certificate and VOMS attributes are delegated to the MyProxy service.
5. The client browser is redirected to the the portal as in figure 2.
6. The portal accesses the grid using the delegated credential.
7. Resources on the grid decide whether or not to grant access, on demand.

CTS in this case. As the MyProxy CA is online, it needs no human intervention to sign certificates. Moreover, the certificates it signs are short-lived, so there is no need for it to revoke them. (In practice, we create an empty revocation list to keep the grid middleware happy.)

The EEC is valid for 7 days and its DN takes the form:
`/DC=uk/DC=ac/DC=ngs/DC=sarongs/CN=`*`pseudo-random ID`*

2.2 Credential Translation Service

The Credential Translation Service (CTS) provides the front end services to grid authentication in this environment. It is an Apache Web server operating only over SSL with a shibboleth protected `cgi-bin`. Within this cgi directory are a number of perl scripts. These scripts make use of functionality in the VOMS::Lite libraries available via CPAN [22]. Each perl script provides a separate piece of the overall functionality of the service, as outlined below.

2.2.1 SARoNGS Login Script

This script provides the most explicit operation of the service (see `HTTPS::GET CTS Page` in figure 4). The figure shows the service with all the network entities involved in the interactive

(portal first) approach. Other modes of operation (e.g. the SARoNGS *Single* script, §2.2.3) do not require such interactive network activity.

1. It obtains a credential from the MyProxyCA.
2. It checks a local VOMSES configuration directory for known VOs.
3. It parses the SAML to obtain IdP scoped attributes, some of which may be understood and used in a VOMS assertion reflecting their role within the IdP's organisation.
4. It renders a web page which presents whether the login has been allowed or denied.

The web page rendering in this case loads some JavaScript functionality. The page presented allows the user to request VOMS attributes they might be entitled to. This is achieved by selecting from known[20] VOs in a drop-down menu or unknown ones by specifying a VOMS server e.g. `voms.gridpp.ac.uk`. When a VO is selected the JavaScript instructs[21] the CTS to use the X.509 certificate obtained above to communicate with the relevant VOMS service. The CTS attempts to obtain the specific credential selected and pass the resulting attribute certificate back to the browser. The resulting attribute certificate is stored as a variable within the JavaScript framework. Any of these operations may succeed or fail for one reason or other. The web page renders success or failure in a number of ways, providing alternatives where appropriate (e.g. if the user is not known to the VOMS server then the web page renders a link to the registration interface, see section 2.2.10). When all choices have been made the user is able to click a continue button which sends all the attribute certificates to the service's *Create* script (§2.2.5). A continue button is present from the beginning should the user wish to proceed without changing any default VO assertions.

By default the user is a member of a VOMS rendering of the federation, a Virtual Organisation named `ukfederation.ngs.ac.uk`. They will have VOMS Group and Role assertions which reflect their eduPersonScopedAffiliation (ePSA). For example the author receives the following:

```
/ukfederation.ngs.ac.uk/manchester.ac.uk/Role=staff
/ukfederation.ngs.ac.uk/manchester.ac.uk/Role=alum
/ukfederation.ngs.ac.uk/manchester.ac.uk/Role=member
/ukfederation.ngs.ac.uk
```

representing */VO/Organisation/Role*, *VO* representing the federation *Organisation* (derived from the scope of the ePSA) and *Role* derived from its value[22].

2.2.2 SARoNGS Logout Script

This function explicitly clears the user's cached credentials from the server. The server regularly cleans up locally stored credentials if there is no current session associated; sessions expire after 5 minutes so as to force the use of fresh SAML assertions. This is necessary for reasons described in section 4.9.

[20] By "known" we mean present in the CTS's *VOMSES* file.

[21] An AJAX XMLHttpRequest asynchronous call is made which expects a plain text reply.

[22] Other VOs will have predefined Groups and Roles, specific to the VO and separate to the attributes received from the IdP.

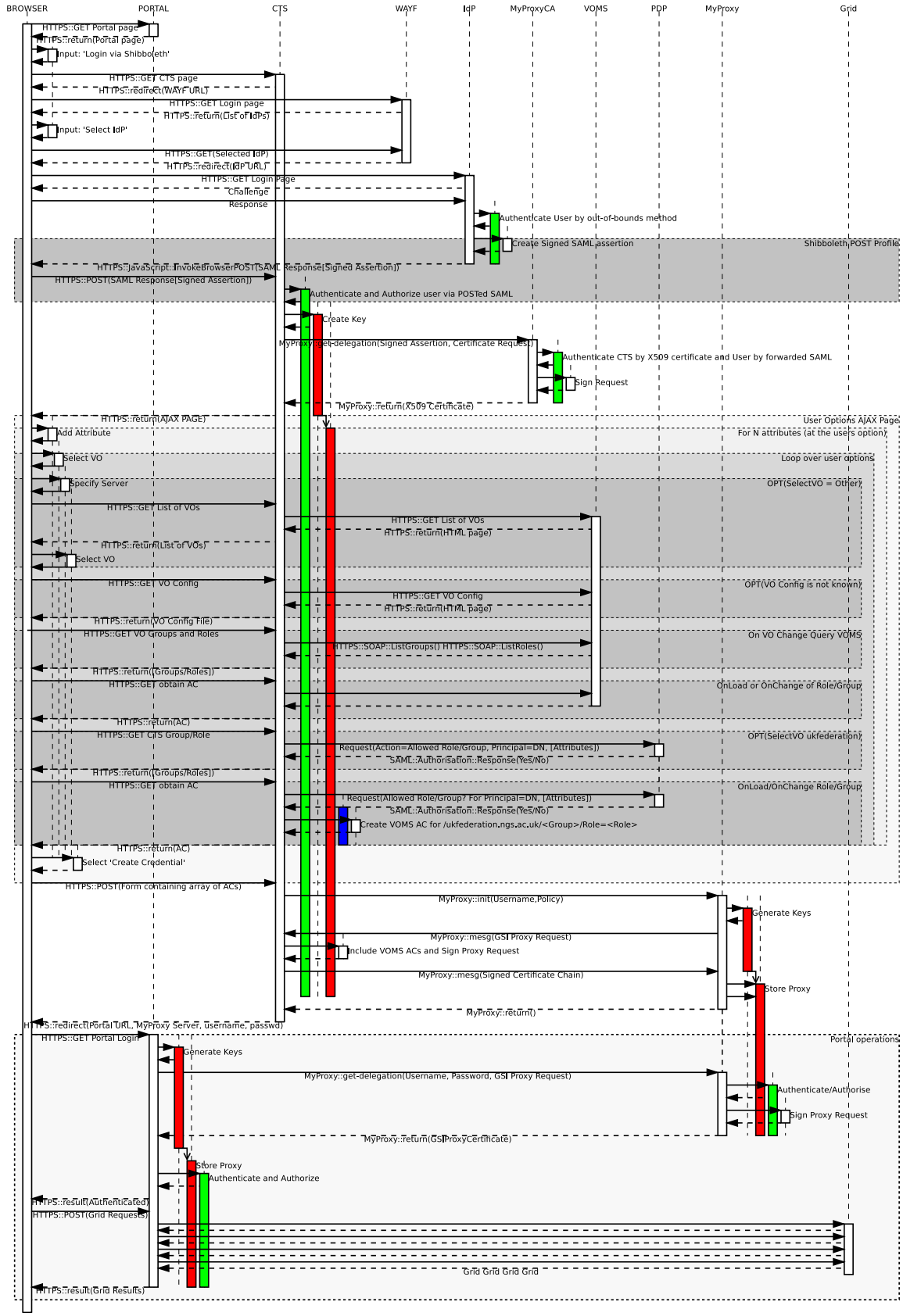

Figure 4: Full Network Activity Schematic

2.2.3 SARoNGS Single Script

This function is designed for portals to make use of. It performs the same operation as Login except all VO credentials are specified in the initial request made by the portal to the CTS. The only page the user will then see after performing their Shibboleth login is a confirmation page where they may cancel the request.

2.2.4 SARoNGS voms-init Script

This function, if called with the correct parameters, returns a VOMS attribute certificate obtained on behalf of the user.
(See `HTTPS::GET Obtain AC` in figure 4).

2.2.5 SARoNGS Create Script

This function performs the delegation of the credential to the world accessible MyProxy server.
(See `HTTPS::Post(Form Containing array of ACs)` in figure 4).

2.2.6 SARoNGS VOMSAdmin Script

This function queries a VO instance on a VOMS server for known Roles and Groups.
(See `HTTPS::GET VO Groups and Roles` in figure 4).

2.2.7 SARoNGS webuiConfig Script

This function queries a VO instance on a VOMS server for its configuration data: the port on which the VO daemon runs, the DN of the certificate that protects it, etc. i.e. the values usually found in a *VOMSES* file.
(See `HTTPS::GET VO Config` in figure 4).

2.2.8 SARoNGS VOMSES Script

This function queries a VOMS service and returns a list of VOs on that server.
(See `HTTPS::GET List of VOs` in figure 4).

2.2.9 SARoNGS PassThrough Script

This function uses the CTS service as an informal HTTP proxy for limited authenticated access to HTTPS pages using the users X.509 credential. This functionality exists to access specific functions of the VOMS service where no obvious APIs or screen scraping methods exist to perform the task.

2.2.10 SARoNGS Registration Script

This function posts a registration request to the VOMS server on behalf of a user. See section 4.4 for more detail.

2.2.11 SARoNGS Confirm Script

This function allows the user to perform the email confirmation step required by the VOMS server. As discussed earlier the VOMS-Admin service assumes that the private key material for the grid credentials is under complete control and ownership of the person and not a service or agent such as this.

2.2.12 SARoNGS RegNGS and ManageNGS scripts and the ngs.ac.uk VO

The ngs.ac.uk VO is an infrastructural level VO designed to serve as a catchall VO[23] for access to resources on the UK's National Grid Service. To date it has served as the primary VO for access to the NGS resources. Membership to this VO requires an additional peer review of each individual's research case and their expected usage of NGS facilities. As such, it has its own management tools beyond of the the scope of the VOMS-Admin service. It does however, rely upon the VOMS services as its means for asserting its users' memberships.

The **RegNGS** function provides the authenticated access to the NGS registration system, the User Accounting System (UAS). The UAS provides accounting information to the VO and the individual. For the user to access this information the **ManageNGS** function is provided.

3. Impact

We understand that each digital asset will have associated with it an access policy and that this is driven by a broad range of requirements on e.g. provenance of data, stability of service, licensing of tools, ownership, copyright, and other legal, operational and policy constraints. We also understand that each individual who wishes to make use of those assets has two main requirements, that the access be as simple as possible and that their personal and research data are not abused. We understand that universities and other agencies called in to identify or otherwise vouch for those individuals often err on the side of conservative attribute release policies[24] or do not have the relevant authority to assert what resource providers are asking them to release[25].

In this paper we hope to demonstrate what we have learned in the process of bridging the gap between specific institutional, community and infrastructure lead identity management environments. There are gaps; there are flaws; there are misconceptions.

Through our experiences we describe short term solutions to the larger problem. By bringing the discrepancies between grid and other communities to light, we hope that all communities can work towards a common understanding of identity and authority.

[23] The possibility for constructing an hierarchical VO representing an array of VOs exists: membership to one VO providing the registration requirements of the next. This would allow for VOs to form and allow access to be granted in a more scalable authorisation infrastructure.

[24] We have come across IdPs which release only that a login has occurred but release no further details about the individual in question.

[25] Experiences gathered during an earlier EPSRC project, "Meeting the design challenges of nanoCMOS technology" [23], where participating institutions IdPs were expected to release project related roles: a delegation of authority which would require formal agreement and process. The best an institute could do, would have been to provide assertions that a researcher was formally assigned to that project. The project ended up by registering users locally on the lead institute's registration system and controlling the release of attributes centrally, resulting in more usernames and passwords for each user to manage.

3.1 Communities

NGS

The NGS started in 2004 as a project to establish a grid environment for all academic research communities in the UK. It started with two general purpose mid–high-end Beowulf clusters, two data storage and processing facilities, and a link into the UK's two high-end supercomputing services CSAR and HPCx. This grid was to evolve by the inclusion of contributed resources from other subscribing institutions, and to date has 29 partner members and affiliates.

The user community is open to include any *bone fide* academic research activity to which a UK based researcher is affiliated. Membership within this community provides the basis to use some or all of the resources provided and affiliated to the NGS. This is achieved by expressing the community as a VO and creating a VOMS instance to represent it in technological terms.

Membership to the NGS VO is peer reviewed and may cover fledgling[26] as well as established research activities within the UK academic community. Once a research activity is established within the grid environment it is expected to form a more formal project or domain specific VO.

GridPP

The GridPP community is the UK particle physics community. Particle physicists are expected to be expert grid users. As such the handling of grid credentials (X.509 certificates and VOMS attributes) is deemed to be within the capability of their researchers. However, within this community there is an outreach activity which provides hands on demonstrations of their research to non-experts see e.g. the VO cernatschool.org. There has recently been an expression of interest in SARoNGS to this effect which has driven the changes to the CA hierarchy described in section 2.1 footnote 19.

OneVRE

OneVRE [24] is a Virtual Research Environment aimed at the Social Science community. It provides a social networking environment to establish communication within research communities. It relied upon SARoNGS to assert VO memberships in order to schedule Access Grid [25] events.

Geospatial

With the advent of the Open Geospatial Consortium [26] standards based online access to geographical data has gained momentum. Part of the SARoNGS project demonstrated secure access to a number of Geospatial services [20].

NeISS

The National eInfrastructure for Social Simulation (NeISS) has made use of SARoNGS infrastructure to provide access for its user base. The users are expected to access resources through the NeISS portal and via the NGS Workload Management System [27] (which is based upon the gLite

[26]Our experience has found that the NGS is able to provide a useful service to early career researchers many of whom would otherwise not be able to access the resources needed to realise their research ideas.

WMS). A document illustrating how SARoNGS can be integrated into portals, with the NeISS portal as the case in point, is available here [28].

4. Experiences

4.1 Using The UK Access Management Federation for Education and Research

The UK Federation provides the initial trust framework within which the Service Providers and Identity Providers operate. Technically the federation provides no legal pathways between IdP and SP, however the federation lay down rules and guidelines for membership [29] and provide the metadata files which practically forms the digital basis of trust between network entities within the federation.

The same rules of membership bind the SARoNGS service to not release attributes about users to third parties. The SARoNGS service could technically release attributes directly through the translation process but decided on the conservative approach of anonymising the attributes received by the service before creating new *anonymised* attributes, so to prevent the release of personal data to other parties. This is down to our interpretation of the rules: if attributes are released by the organisation which hosts the Service Provider, has it then broken the rules, or can the global (in the IGTF case, or in this case national) grid be seen as a single service provided by the organisation? The latter doesn't seem plausible, so we end up not releasing attributes except in encrypted (hashed) form.

4.2 The Use of PKI

Grid infrastructures chose PKI to mitigate against having to maintain high-availability services for authentication, such as Kerberos, see [2]. This requires a trust model to be in place where the service provider can rely upon the PKI Hierarchy to be able to identify any user accessing their system who possesses an authentic grid credential (their X.509 certificate). The hierarchy is constructed with a series of authorities which delegate the responsibility of registering an individual or service identity to cryptographic key material, bind these individuals to policies associated with the management of that key material, and label them with a unique identifier (their Distinguished Name). Finally the IGTF takes on the role of measuring these authorities operational procedures and service level definitions against reference policy templates and each other, providing a *rubber stamp* of trustworthiness.

Usually within a PKI the certificate authority performs registration and this will be based upon a personal interview with a Registration Authority (a role within the organisation of the CA). SARoNGS is different in that it takes an institute's already established registration processes and uses that as the basis for registration within the PKI, thus substituting a 1–7 day registration process with a nearly instantaneous exchange of cryptographic information[27].

[27]This is similar to the processes within the IGTF MICS profile [30] which, briefly, allows certificates to be issued based on institutional identity; the difference being that we are working here with fewer attributes and our IdPs cannot in general satisfy the auditing requirements imposed by IGTF (see also §4.7).

4.3 Use of Signed Attribute Statements

SAML is capable of carrying digitally signed attribute statements[28]. In SARoNGS these attribute statements are used on-the-fly to form the basis of user registration, avoiding yet another strong identification process. The trust between the IdP and the online CA is established by the requirement to have signed attribute statements, which are recorded by the SARoNGS service. The institute thus provides a digitally signed document asserting the affiliation of a user and a traceable identifier to that individual. This, in turn, allows the CA to issue a credential to the CTS knowing that there was an authentic login to the CTS a short time before.

If the online CA did not require the credential to be signed and verified the CTS could manufacture bogus SAML assertions and have the CA issue any X.509 certificate at will. This requirement for a signed assertion therefore significantly reduces the risk of issuing certificates to a compromised CTS service should that ever come to pass, the CTS being one part in an abstract two factor authentication process[29].

However, XML signing is notoriously difficult. During operation of the SARoNGS service the Shibboleth implementation used was discovered to manipulate the SAML assertions before it was passed on to the environment, occasionally inserting extra XML Namespace headers. This would alter the canonicalised XML, and thus the message digest would be evaluated incorrectly and the signature verification would ultimately fail. The SARoNGS service was forced to catch these extra headers and rectify them at the application layer before sending them on to the CA.

4.4 VO Registration

Registering with a VOMS hosted VO under normal circumstances requires that the user has their grid certificate installed in their browser. A user would visit a VOMS-Admin instance (for example https://voms.ngs.ac.uk:8443/voms/open.vo.ngs.ac.uk[30] and follow registration instructions. During this process they would have to fill in a form asserting (unverified) information about themselves: *Name, email Address, Address, Phone Number*. A verification email[31] would then be sent to the user containing a secure link which they must follow (again with their grid-certificate in their browser). The SARoNGS service, however, does not release the end entity certificate that would usually be installed in the user's browser and so *it* must perform the tasks above on behalf of the user.

If a user attempts to obtain a VOMS assertion to which they are not entitled, the CTS will notify the user. That notification will contain a link to register with the VO through the CTS. A secure connection to the CTS is already established through the Shibboleth federated identity mechanisms. The link provided will prompt the user for the usual VO registration data and the CTS will POST this to the VOMS-Admin service over the usual mutually authenticated HTTPS using the user's stored EEC.

[28] SAML2 also provides encryption.

[29] *Something you have:* a time limited assertion about the user, targeted at the CTS; *Something you know*: the CTS's private key.

[30] The open.vo.ngs.ac.uk is a test VO provided by the NGS to allows any user to register, administer and test interactions with a VOMS server.

[31] This would not be necessary if the user's email address was expressed as an attribute within their EEC (which is often the case), and VOMS respected this assertion.

It is not straight forward to perform the email registration, however. The user would usually receive an email after initial registration with a link to follow to confirm that their email address is correct and valid. If the user follows that link, the VOMS-Admin service will be unable to process the request as it won't be sent over the mutually authenticated HTTPS. Instead, for SARoNGS, the user must pass that link to the CTS who can complete the registration request.

The user is made aware of this at the time of registration through the CTS web interface. The CTS also issues an email with counter instructions.

4.5 VO Administration

For the VO administrator when approving a VO membership they will have two pieces of authentic information about the user: their CommonName (derived from their certificate's Distinguished Name) and their email address. In the case of the SARoNGS work-flow the CommonName attribute available to the administrator will be pseudo-anonymous[32]. As such the only handle of value for identifying the individual will be their email address. If this is not recognisable in itself, it does at least provide a point of communication.

4.6 Debugging

The nature of federated identity management and the use of various Identity Provider middleware versions etc. results in a difficult operational environment. Should a system not work for one particular person or organisation, it is often not possible for a developer to reproduce the effect, given that they are not usually registered within that IdP's institute. During the development phase of SARoNGS a number of issued arose where debugging required the users to be present to perform occasional login operations and report.

Any iterative approach to debugging rapidly becomes very difficult.

4.7 The Grid and Accreditation within Grids

To drive this solution into full production and achieve the main goal of SARoNGS, the certificates that it produces need to be able to be validated and the SARoNGS CA hierarchy needs to be trusted on the grid resources. It has been difficult to have resources trust certificates outside of the IGTF, that being the *de facto* standard distribution for grid certification authorities' issuing certificates. Unfortunately it is not possible to register the SARoNGS CA with the IGTF under any current profile at this time for the following reasons:

- No reasonable representation of the users name is published: we don't have a *CommonName* attribute to use[33].
- By the Rules of Membership (RoM) [29], attributes may be used only for representation and access control, or may be aggregated statistically, uses which may not be sufficient for the grid (where DNs are currently public data).

[32] In reality, even a well defined CommonName is not strictly enough to identify an individual without some other out-of-band identification mechanism.

[33] It is argued by some that this is in fact not necessary, as the identifier for the individual is the Distinguished Name and thus security alerts using that DN should be sufficient to satisfy security policy; it is argued by others that it is in deed requirement so that softer security, incident response procedures can be implemented where an exact match to a DN does not work e.g. in two separate domains using different identity technologies.

- Attributes may not be passed to third parties – this is usually interpreted as "must not leave the legal entity (organisation) that hosts the service provider (see also §4.1."[34]
- Accountability in authentication is optional (known as "section 6" in the RoM). This is a record of what links a person to their IdP identity.
- Even for accountable institutions, the principal authentication token (the ePTID) can be recycled.
- No guarantee that users were checked by approved photo id (i.e., that IGTF requirements for identity vetting were followed) – although each IdP is required to be willing to provide documentation of its processes, these will have to be requested from each and reviewed.
- IdPs never release personal data; in particular, they cannot be audited.

4.8 Roots of Trust for the Service

One well known difference between commercial certificate authorities and those generally employed within grid computing infrastructures is how they are installed. Commercial CAs are generally shipped with the operating system or the browser. This is seen as a benefit for the majority of users who prefer not to have to install and manage trust anchors on their systems themselves.

Systems provided for grid infrastructures often choose, and in some cases are required, to have their identities certified by a CA from the grid domain (usually by an IGTF accredited CA). For the user to use these services they usually have to install and manage the grid trust anchors themselves. VOMS services are a case in point. They use their host certificate for three distinct purposes:

- As the Attribute Authority certificate, used to sign VOMS assertions,
- As the server certificate for the VOMS-Admin web interface,
- As the service certificate for the VOMS services (which are for example contacted by the user during a `voms-proxy-init`).

and therefore this certificate must be appropriate for use within the grid paradigm.

For historical reasons there are tools which query VOMS services to obtain its grid certificate. So to satisfy both ease of use for users, and applications which query VOMS services[35] all CTS services are available on two ports: Port 443 the usual web HTTPS port secured with a certificate issued by a commercial CA and port 8443 a second HTTPS port secured by a grid domain certificate. This setup is known as *port based virtual hosting* and allows browser access to proceed without security warnings, while providing a service that is compatible with the grid's security federation.

4.9 AJAX and Shibboleth

Shibboleth employs a protocol which sits in both the application layer and the HTTP layer. During the login process the user is required to authenticate to a third party Identity Provider, usually via a web-form. AJAX employs HTTP request methods which are for the consumption

[34]The SARoNGS service sidesteps this issue by mangling one attribute before releasing it as an identity handle and making an authorisation decision based upon a collection of other attributes before releasing VOMS Fully Qualified Attribute Name (FQAN) attributes.

[35]The SARoNGS also provides a bespoke VOMS Attribute Authority service.

of the JavaScript environment within the web page and not the user. If the Shibboleth security context expires, any request that attempts to access a resource within the protected Shibboleth area will be redirected to a login page or WAYF: a result that is difficult to handle gracefully within the JavaScript framework. To mitigate against this, the Login script (§2.2.1), which employs AJAX, will auto-refresh within the context's lifetime after removing all CTS related Shibboleth context cookies. This will force the user to login again and the page will therefore never be in a position where the AJAX HTTP calls are denied access due to an expired session.

4.10 Too Many Clicks

The SARoNGS Login script, as described in section 2.2.1, presents all option available for the user in one web page login. The feedback from the community is that this is too complex. To address this the SARoNGS Single script (§2.2.3) was created. This moves the complexity to the communities' portals to make choices or present complexities in a way that is more appropriate to their user base. The SARoNGS Simple script results in a process where the user clicks to login, clicks to choose their home institute, logs into their home institute, and clicks to allow the CTS to release credentials.

Even this has been deemed too difficult by some communities and as part of the CONSENT [31] project — a short community focused follow-on project for SARoNGS — steps are being made to address this. The CONSENT project aims to deliver community focused login environments where the last step can be all but removed[36]. Two approaches are currently under review

- A branding based exercise where the look and feel of the community portal is brought in house to the CTS, so the user does not appear to lose the presentation layout they expect,
- An iframe approach where the CTS login pages are presented within the users portal.

While these do not reduce the number of clicks, they may appease the user community by not changing the style associated with their portal experience.

5. Conclusions

The UK houses two separate authentication authorisation infrastructures for academic research, one relatively easy-to-use system based on the relationship between individuals and their Universities; the other more complex, based upon the individual's personal credentials under their own management and the out of band subscription to Virtual Organisations.

In reality these are complex systems which don't align themselves very well. We have bridged the gap at least technically, and aimed to work within the limitations imposed by policy requirements, both nationally and internationally. Where necessary plugged holes to produce a production service which presents the simpler Shibboleth authentication environment and uses this to issue the more complex grid credentials supplying them to the portals to use on behalf and at the request of the user.

We have learned much in the process and have solved many issues allowing researchers in the UK access to UK based grid resources. On a larger scale we have identified sociopolitical barriers

[36]The click to release credentials is part of a necessary authorisation step that the user must agree to. Without this step, it is possible for the portals to abuse the CTS and obtain credentials that the user did not wish to bestow.

to uptake, which we hope to solve through EGI, federation and standards activities. In other future work we hope to embrace the efforts of the Moonshot project [32].

References

[1] Wang, X.D., Jones, M., Jensen, J., Richards, A., Wallom, D., Tiejun, M., Frank, R., Spence, D., Young, S. et al. "Shibboleth Access for Resources on the National Grid Service (SARoNGS)". *Journal of Information Assurance and Security* 5, 3 (2010) pages 293–300

[2] Foster, I., Kesselman, C., Tsudik, G. and Tuecke, S. A Security Architecture for Computational Grids. *ACM Conference on Computers and Security* (1998) pages 83–91

[3] UK Access Management Federation for Education and Research. Online http://www.ukfederation.org.uk/ [Accessed 15/04/2012]

[4] Kelsey, D. (On behalf of the EGI Security Policy Group) "VO Portal Policy" EGI-doc-80-v7 (08 March 2011) Online https://documents.egi.eu/document/80 [Accessed 15/04/2012]

[5] Foster, I., Kesselman, C., Tuecke, S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *Int. J. High Perform. Comput. Appl.* 15, 3 (August 2001) pages 200–222

[6] David Groep (Ed.) "International Grid Trust Federation" IGTF-Federation-20051005-1.doc (25 June 2008) Online http://www.igtf.net/new-doc/IGTF-Federation-20051005-1-igtf.pdf [Accessed 15/04/2012]

[7] Grid and Shib: Investigators meeting at the 16th Global Grid Forum. Online http://www.ogf.org/gf/event_schedule/?id=213 [Accessed 15/04/2012]

[8] JISC Middleware technology development Programme: ShibGrid project. Online http://www.jisc.ac.uk/whatwedo/programmes/middleware/shibgrid.aspx [Accessed 15/04/2012]

[9] JISC Middleware technology development Programme: SHEBANGS project. Online http://www.jisc.ac.uk/whatwedo/programmes/middleware/shebangs.aspx [Accessed 15/04/2012]

[10] Tuecke, S., Welch, V., Engert, D., Pearlman, L., and M. Thompson, "Internet X.509 Public Key Infrastructure (PKI) Proxy Certificate Profile", RFC 3820 (June 2004) (http://www.ietf.org/rfc/rfc3820.txt [Accessed 15/04/2012])

[11] J. Basney, M. Humphrey, and V. Welch. The MyProxy Online Credential Repository. Software: Practice and Experience, Volume 35, Issue 9, July 2005, pages 801–816

[12] National Institute for Standards and Technology, Special Publication 800-63: Electronic Authentication Guideline v1.0.2, April 2006, Online http://csrc.nist.gov/publications/nistpubs/800-63/SP800-63V1_0_2.pdf [Accessed 15/04/2012]

[13] Alfieri, R et.al. Managing Dynamic User Communities in a Grid of Autonomous Resources. Corr cs.DC/0306004 Online http://arxiv.org/abs/cs.DC/0306004 [Accessed 15/04/2012]

[14] Ciaschini, V., Venturi, V., Ceccanti, A. "The VOMS Attribute Certificate Format", GFD.182 (August 2011) (http://www.ogf.org/documents/GFD.182.pdf [Accessed 15/04/2012])

[15] Chadwick, D.W., Otenko, O. The PERMIS X.509 role based privilege management infrastructure, *Fut. Gen. Comput. Stsy. 19(2)*, Elsevier Science Publishers B.V., 2003, pages 277–289

[16] Maler, E., Mishra, P., Philpott, R. (Eds.) "Assertions and Protocol for the OASIS Security Assertion Markup Language (SAML) V1.1.", oasis-sstc-saml-core-1.1 (2 September 2003) Online https://www.oasis-open.org/committees/download.php/3406/oasis-sstc-saml-core-1.1.pdf [Accessed 15/04/2012]

[17] "Security Requirements for Cryptographic Modules", *Nat'l Inst. of Standards and Technology*, FIPS PUB 140–2 (25/05/2001) Online http://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf

[18] Basney, J. "MyProxy Protocol", GFD.54 (November 2005) (http://www.gridforum.org/documents/GFD.54.pdf [Accessed 15/04/2012])

[19] The NGS Portal. Online https://portal.ngs.ac.uk/ [Accessed 15/04/2012]

[20] Millin, G. and Kitmitto, K. A Review of the Landmap Service New Data Acquisition and Data Delivery Infrastructure at the Mimas National Data Centre. PROCEEDINGS OF THE REMOTE SENSING AND PHOTOGRAMMETRY SOCIETY CONFERENCE 2008 "Measuring change in the Earth system". University of Exeter, 15–17 September 2008. (http://www.landmap.ac.uk/images/stories/support/publication/papers/rspsoc_millin_kitmitto_2008.pdf [Accessed 15/04/2012])

[21] Hazelton, K. (Ed.) "EduPerson Object Class Specification" internet2-mace-dir-eduPerson-200604a (May 15, 2007) Online http://middleware.internet2.edu/eduperson/docs/internet2-mace-dir-eduperson-200604.html [Accessed 15/04/2012]

[22] VOMS::Lite. CPAN Repository. Online http://search.cpan.org/~mikej/VOMS-Lite/ [Accessed 15/04/2012]

[23] Sinnott, R.O., Stewart, G., Asenov, A., Millar, C., Reid, D., Roy, G., Roy, S., Davenhall, C., Harbulot, B., and Jones, M. "E-infrastructure support for nanoCMOS device and circuit simulations", *Proceedings of the Conference on Parallel and Distributed Computing and Networks* Hamza, M.H. (Ed.) (February 2010) ACTA Press

[24] Turner, M., Schiebeck, T., Poschen, M., Jones. M., Rowley, A. Creating a Secure Distribution Cross-Portlet System: Experiences from OneVRE Phil. Trans. R. Soc. A pre-submission for 2012 print

[25] Stevens, R. "Access Grid: Enabling Group Oriented Collaboration on the Grid", *The Grid: Blueprint for a New Computing Infrastructure* C. Kesselman (Ed.) (2003) Morgan Kaufmann

[26] The Open Geospatial Consortium Online http://www.opengeospatial.org/ [Accessed 15/04/2012]

[27] The NGS UI-WMS Resource Broker. Online http://www.ngs.ac.uk/ui-wms [Accessed 15/04/2012]

[28] Watt, J. Integrating SARoNGS with Portal Infrastructures. Online http://www.escience-etf.ac.uk/documents/SARoNGS-Documentation.pdf [Accessed 15/04/2012]

[29] UK Access Management Federation for Education and Research Rules of Membership. August 2011. Online http://www.ukfederation.org.uk/library/uploads/Documents/rules-of-membership.pdf [Accessed 15/04/2012]

[30] Murray, M. "Profile for Member Integrated X.509 Credential Services with Secured Infrastructure", IGTF-AP-MICS-1.1, (May 2009) (http://www.tagpma.org/files/MICS_AP_FINAL_1.1.pdf)

[31] The JISC Access and Identity Management Programme: Communities On the NGS via SARoNGS ENabled Trust project website. Online http://www.jisc.ac.uk/whatwedo/programmes/di_directions/accessandidentitymanagement/consent.aspx [Accessed 15/04/2012]

[32] The Moonshot project. Online http://www.project-moonshot.org/ [Accessed 15/04/2012]