Single Sign-on and Authorization for Dynamic Virtual Organizations

R.O. Sinnott, O. Ajayi, A.J. Stell, J. Watt, J. Jiang
National e-Science Centre
University of Glasgow
Glasgow, Scotland
ros@dcs.gla.ac.uk

The vision of the Grid is to support the dynamic establishment and subsequent management of virtual organizations (VO). To achieve this presents many challenges for the Grid community with perhaps the greatest one being security. Whilst Public Key Infrastructures (PKI) provide a form of single sign-on through recognition of trusted certification authorities, they have numerous limitations. The Internet2 Shibboleth architecture and protocols provide an enabling technology overcoming some of the issues with PKIs however Shibboleth too suffers from various limitations that make its application for dynamic VO establishment and management difficult. In this paper we explore the limitations of PKIs and Shibboleth and present an infrastructure that incorporates single sign-on with advanced authorization of federated security infrastructures and yet is seamless and targeted to the needs of end users. We explore this infrastructure through an educational case study at the National e-Science Centre (NeSC) at the University of Glasgow and Edinburgh.

1. INTRODUCTION

The vision of the Grid [7] is to support the *dynamic* establishment and subsequent management of virtual organizations (VOs). The term *dynamic* is italicized here as it could be argued that this is what distinguishes Grid infrastructures from other large scale distributed systems. With no prior detailed agreements in place, it should be possible to create a VO which will allow collections of individuals and/or institutions to *securely* share resources whether these resources are data sets, data archives, computational resources, services or more specialized equipment. A key element of this vision is the notion of *single sign-on* where a single set of user security credentials are sufficient to allow access to a multitude of federated resources across the VO.

Perhaps the greatest challenge in realizing this dynamic model is security. Sites wishing to potentially form a VO need to be aware of the consequences of establishing such collaborations. It is the case in computer security that the weakest link rule applies; this fact is magnified by Grid infrastructures due to their openness. Highly secure multi-million pound compute facilities can be compromised by inadequately secured remote laptops. Rigorous security procedures at one site can be made redundant through inadequate procedures at another collaborating site. This problem is exacerbated by the predominant Public Key Infrastructure (PKI) [11] authentication-only based security models prevalent across most high performance computing related Grid infrastructures today, where establishment of user identity is the primary security focus (and not on restricting what the user is allowed to do on

the given resource). With the move of the Grid community to more security focused areas such as the health domain, this authentication-only security model is unrealistic and does not lend itself to the adoption of Grid technology. Considerable progress has been made in developing advanced security infrastructures that are well integrated into Grid middleware [8,20]. However the challenge remains how to establish a VO in a dynamic manner where sets of fine grained distributed security authorization policies defining what end users are allowed to access/use on local institutional resources can be supported across multiple independent institutions.

One common approach to solve this is through the establishment of *federations* which can be considered as groups of organizations which agree to adopt common policies and technical standards to provide a common infrastructure for managing access to resources and services in a uniform way. The Internet2 Shibboleth architecture and protocols [18,19] have been developed to support the establishment of federations where devolved authentication and security attribute release across multiple independent institutions is supported. Through Shibboleth, authentication at a home institution Identity Provider (IdP) by a user can in principle support single sign-on across a federated VO where security attributes and assertions are released which can subsequently be used by service providers (SP) to make authorization decisions. This model of single sign-on lends itself to advanced authorization in more security focused VOs, but requires detailed negotiation of security attributes to be defined *a priori*. This pre-agreed and potentially detailed negotiation and agreements goes somewhat against the true vision of the Grid where dynamic VOs can be established and managed "on-the-fly", and where new agreements and policies can be added as new institutions, new resources and users are brought together for potentially short time periods.

In this paper we outline a novel solution prototyped within the UK JISC Dynamic Virtual Organizations for e-Science Education (DyVOSE) project [5] that, using a basic institutional trust relationship between sites supports single-sign combined with advanced authorization of federated security infrastructures based upon delegation of authority. We explore this infrastructure in an educational setting through a programming assignment set as part of the Grid Computing module part of the advanced MSc at the University of Glasgow.

2. EXISTING GRID SECURITY LIMITATIONS

Grid security is still predominantly based around PKIs to support authentication, i.e. the validation of the identity of a given user requesting access to a given resource. The simplest PKI involves a single Certification Authority (CA) which is trusted by all users and resource providers. With this model, users only accept certificates (signed by the CA which associate the users private key with their public key) and certificate revocation lists issued by this CA. This model makes certificate path analysis easy since there is a single step from a certificate to the CA who issued it.

Other more complex PKI architectures also exist. For example, users may keep a host of trusted CAs. However, issues such as how to tell trustworthy one from untrustworthy one arise. Hierarchical PKIs where there are chains of trust between the CA, sub-ordinate CAs and users may also exist. This model allows limiting the damage caused by compromised subordinate CAs. Thus if a subordinate CA is compromised then only the certificates issued by them (or their subordinate CAs) need to be revoked. Other more complex architectures exist again, such as meshes of

PKIs where trust relationships (webs of trust) are established on a peer-peer basis. This model often requires bridging solutions [12,15] between CAs and results in certificate paths that are harder to establish – potentially containing loops.

The main benefit and reason for the widespread acceptance of PKIs within the Grid community is their support for *single sign-on*. Since all Grid sites in the UK trust the central CA at Rutherford Appleton Laboratories (RAL) [23], a user in possession of an X.509 certificate issued by this CA can send jobs to all sites, or rather to all sites where a user has requested and been granted access. Typically with Globus based solutions gatekeepers are used to ensure that signed Grid requests are valid, i.e. from known collaborators. This is manifest through the Distinguished Name (DN) of the requestor being in a locally managed access control list (ACL) grid *mapfile* which typically maps DNs to local user accounts. These ACLs are typically manually updated and managed based upon individual user requests. The dynamicity of this manual approach is not conducive to the Grid-idea for dynamically establishing new, potentially short term VOs. Instead users have to statically have their DNs registered at collaborating sites which have previously made available/allocated local accounts. Once the Grid scales to the wider research and academic communities (as opposed to the current focus on the "Grid" community) where many millions of users[1] exist this centralized model of certification is likely to have scalability issues.

The process of acquiring an X509 certificate itself is off-putting for many of the less-IT focused research community since it requires them to convert the certificate to appropriate formats understandable by Grid middleware, e.g. through running cryptic (in the confusing sense!) *openssl* commands [13]. This problem is further exacerbated since *openssl* is not commonly available on platforms such as Windows and requires separate software to be installed. Once in the appropriate Grid format, users are then obliged to remember necessarily strong 16-character passwords for their certificates with the recommendation to use upper and lower case alphanumeric characters. The temptation to write down such passwords is apparent and an immediate and obvious potential security weakness.

The fundamental issue with PKIs for Grid security however, is trust. Sites trust their users, the CA and other sites. If the trust between any of these is broken, then the impact can be severe, especially since users are currently free to compile and run arbitrary code. With the now global PKI and associated recognition of international CAs through efforts such as the International Global Trust Federation (www.gridpma.com), this basic trust model is naïve. For this reason, Grids have been seen as at best something to be considered separately from existing compute infrastructures or at worst as a potential threat to those infrastructures.

3. SINGLE SIGN ON AND ADVANCED GRID SECURITY FOR ***STATIC*** VOs

Numerous technological solutions have been put forward looking towards providing various enhanced Grid security models and solutions such as CAS [14], GSI [9], PERMIS [2] and VOMS [1]. Examples of how these compare to one another is

[1] There are currently over 3 million Athens accounts across UK academia from over 2,000 organizations. To put this into context there are approximately 3500 UK e-Science certificates issued by the UK e-Science CA that are currently valid across the UK.

described in [21]. Recent developments in Grid standardization [8] and associated implementations [3] have shown, however, how finer grained models of security can be achieved supporting authorization closely integrated with Grid solutions.

Role Based Access Control (RBAC) based solutions represent one of the more scalable solutions for advanced authorization infrastructures [permis]. Such systems allow for definition of roles which are typically associated with given privileges on a system and as such, are less susceptible to change than individual user identities. The roles themselves are assigned to subjects (users) by issuing them with an X.509 attribute certificate (AC) [2]. These roles and ACs can in turn be used to form the security policies for a given site. Systems such as PERMIS allow for the expression of digitally signed (and hence tamper proof) security policies based upon triplets comprised of <*Role, Target, Action*>. A local authority – the Source of Authority (SoA) will specify policies based upon institutional roles, institutional resources (targets) and actions that can be performed on those resources. Once defined, these policies can be used to ensure that only users with appropriate roles (privileges) can access certain services or data resources and perform certain actions. It has been shown [22] how such infrastructures can be defined and used as the basis for limiting access to Grid resources and data sets. Such systems predominantly work at the local authorization level, i.e. the policies apply to the local site only. With Grid based inter-institutional VOs this model of security is not the norm and collective understanding of inter-institutional security infrastructures is needed.

Supporting multiple attribute authorities is something that the Internet2 community has focused on explicitly in the Shibboleth architecture and protocols [18,19]. The UK academic community is currently in the process of deploying Shibboleth technologies (http://shibboleth.internet2.edu/) to support local (existing) methods of authentication for remote login to resources. Through this model, sites are expected to trust remote security infrastructures for example in establishing the identity of users (authentication) and their associated privileges (authorization). To support this, the Shibboleth architecture and associated protocols identify several key components that should be supported including federations, Identity Providers (aka origins), Service Providers (aka targets) and optionally Where Are You From (WAYF) services. Through these components, end users will have single usernames and passwords from their home institutions which will provide for seamless access to a range of resources at collaborating institutions and service providers. Local security policies at service provider sites can then be used to restrict (authorize) what resources authenticated users are allowed access to.

To understand the impact of Shibboleth technologies on Grid security it is first necessary to have an appreciation of the interactions that typically arise with Shibboleth. When a user attempts to access a Shibboleth protected service or Service Provider (SP) more generally, they are typically redirected to a WAYF server that asks the user to pick their home Identity Provider (IdP) from a list of known and trusted sites. The service provider site has a *pre-established trust relationship* with each home site, and trusts the home site to authenticate its users properly.

After the user has picked their home site, their browser is redirected to their site's authentication server, e.g. an LDAP repository, and the user is invited to log in. After successful authentication, the home site redirects the user back to the SP and the message carries a digitally signed SAML [17] authentication assertion message from the home site, asserting that the user has been successfully

authenticated (or not!) by a particular means. The actual authentication mechanism used is specific to the IdP. If the digital signature on the SAML authentication assertion is verified and the user has successfully authenticated themselves at their home site, then the SP has a trusted message providing it with a temporary pseudonym for the user (the handle), the location of the attribute authority at the IdP site and the service provider URL that the user was previously trying to access. The resource site then returns the handle to the IdP's attribute authority in a SAML attribute query message and is returned a signed SAML attribute assertion message. The Shibboleth trust model is that the target site trusts the IdP to manage each user's attributes correctly, in whatever way it wishes. So the returned SAML attribute assertion message, digitally signed by the origin, provides proof to the target that the authenticated user does have these attributes. We note that later versions of the Shibboleth specification have introduced a performance improvement over the earlier versions, by allowing the initial digitally signed SAML message to contain the user's attributes as well as the authentication assertion. Thus the two stages of authentication and attribute retrieval can be combined.

This security model offers several direct benefits over PKIs for dynamic establishment of VOs in that users are no longer trusted to manage their X509 certificates and remember complex passwords. Instead institutions within a federation have a degree of trust with one another. Sites/IdPs and SPs are still autonomous and are able to decide for themselves whether the provided attributes are sufficient for access to the resources and which attributes they are prepared to release to which SP. Another key benefit of Shibboleth for VO establishment and management is that users are only required to remember their own usernames and passwords at their home institutions.

Provided a common understanding of the roles and security attributes across the sites comprising the federation exists, single sign on can be achieved. Thus if a SP trusts a given site for authenticating a user requesting access to its own resource, and also an agreement on the attributes which are to be exchanged between the sites exists, then the SP can authorize/restrict access to its resources from those sites that are within the correct federation and that provide the necessary attributes and their values needed to give access to the resource. Within the UK a single federation is being proposed (www.sdss.ac.uk) and a small set of security attributes based upon a subset of the eduPerson specification is being adopted [16]. These attributes include *eduPersonScopedAffiliation* which indicates the user's relationship (e.g., staff, student, etc.) within their home institution; *eduPersonTargetedID* which is needed when an SP is presented with an anonymous assertion only as provided by *eduPersonScopedAffiliation*; *eduPersonTargetedID* attribute which provides a persistent user pseudonym; *eduPersonPrincipalName* which is used where a persistent user identifier, consistent across different services is needed, and *eduPersonEntitlement* which enables an institution to assert that a user satisfies an additional set of specific conditions that apply for access to a particular resource. A user may possess different values of the *eduPersonEntitlement* attribute relevant to different resources.

One key aspect of the UK federation which helps to support single sign-on across numerous resources is the facility to maintain session information. Thus in accessing their IdP, the user is able to specify whether the WAYF should remember them for the duration of the session, for a week or not at all. In accessing subsequent

Shibboleth protected services, the WAYF will automatically recognize which IdP the users are from and redirect them accordingly.

Proof of concept systems demonstrating how Shibboleth based access to Grid resources has been achieved is described in [25]. However Shibboleth by its very nature is much more static than the true vision of the Grid, where VOs can be dynamically established linking disparate computational and data resources at run time. Instead Shibboleth requires agreed sets of attributes that have been negotiated between sites. What is needed instead is a more dynamic way in which security attributes associated with a VO can be established and accepted across a given federation.

4. SINGLE SIGN-ON AND ADVANCED GRID SECURITY FOR ***DYNAMIC*** VOs

The definition of detailed policies for access to and usage of multiple site resources will face scalability issues for large scale Grid infrastructures where many different users, services and resources exist. This is further compounded when new users join, leave, new resources are added and removed etc. Having a single SoA to manage a security infrastructure at a given site is not realistic for large scale, evolving Grid infrastructures. Ideally, it should be possible to *delegate* the privilege for others including potentially those at other trusted sites to issue ACs which will be recognized locally. This is especially the case when complex or short lived dynamic VOs are to be established and managed. To address this, the DyVOSE project has prototyped a delegation issuing service (DIS) [3] as shown in Figure 1.

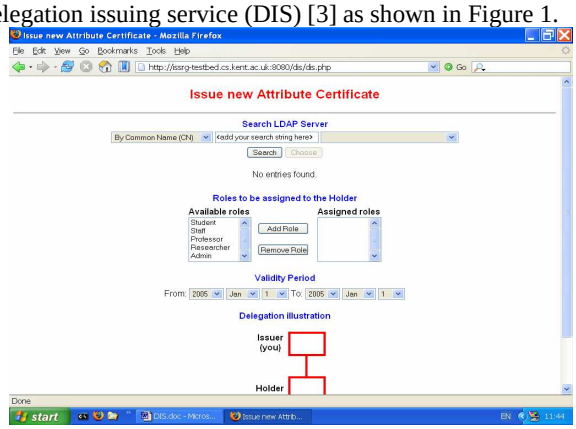

Figure 1: DyVOSE Delegation Issuing Service

The DIS is a web service that can issue ACs on behalf of a SoA. In a traditional PMI model a SoA that issues roles or privileges to users must have a PKI key pair. This restriction implies that the SoA is responsible for all privilege allocations within at its own site. Ideally a SoA, e.g. a systems-administrator, would like to be able to delegate the privilege to issue new roles to other trusted entities/people, e.g. to a local researcher wishing to establish a particular VO, or potentially to a remote but trusted entity. The DIS service itself does not require that delegated/trusted entities are required to hold a PKI key pair in order to issue ACs to their subordinates however the SoA will by definition restrict the roles that its subordinate authority will be able to issue. The DIS service also allows delegated entities to also delegate

privileges to others. To minimize the potential security risks that might arise through this, subordinate authorities will always have lower privilege than their superiors.

Through the delegation of authority capabilities offered by the DIS service, sites wishing to establish VOs dynamically are able to create attribute certificates associated with the particular demands of the give VO. Once defined, users wishing to access resources across multiple institutions are able to use the single sign-on capabilities of Shibboleth to authenticate themselves at their home site, and have these attributes (which have been dynamically created) to be used by SPs to make subsequent authorization decisions. Through this, dynamic VOs can be established where fine grained authorization policies are created based upon attributes specific to the security of the VO and created by privileged members of the VO. Subsequent access to Grid resources across the VO can, through Shibboleth, be based upon the appropriate attributes being defined and subsequently delivered for authorization decisions to be made to VO resources.

To explore the capabilities of the DIS service for dynamic attribute creation and their usage for subsequent single sign-on through Shibboleth to access and use dynamically established VO resources, we have explored this technology within the advanced MSc Grid Computing module at the University of Glasgow.

4.1 Case Study

The Grid Computing module at the University of Glasgow required the advanced MSc students to undertake a large scale programming assignment. This assignment was focused on exploring latest developments in Grid middleware such as Globus and Condor [4], and exploring fine grained security infrastructures. Specifically, the students were required to implement a Globus-based bioinformatics application (BLAST) which was to run across a Condor pool. The application required them, in the first instance to develop a client to access a remote Grid service (*BlastData*) in Edinburgh University which was protected by the PERMIS authorization infrastructure and return the appropriate sequence data. This service and the associated security policy was developed and deployed in advance for the students. The students were split into two groups: *groupA* and *groupB*. These groupings (roles) were then used by the *BlastData* service and its security infrastructure to enforce/restrict access to the data accessible. The data itself was nucleotide or protein sequence data sets depending on the role (group) the students were in.

Once the data was returned the students were expected to use this as input to their own Globus based BLAST service which would run across the Condor pool. This service was also PERMIS protected with the policy such that only members of their team could invoke the service, i.e. people with their role. Diagrammatically the assignment and associated infrastructure is given in Figure 2.

In the infrastructure the *Glasgow SoA* used the Edinburgh DIS service to issue attributes within the Edinburgh PMI for roles needed across the VO, i.e. they were delegated the privilege by the *Edinburgh SoA* to create roles within the Edinburgh role hierarchy. Through creation of a VO specific role, e.g. *externalStudent* within the Edinburgh policy via DIS and mapping of the DNs of Glasgow students to this role, Glasgow students have subsequently been able to access and return the appropriate sequence data sets for input to the BLAST service. Through the hierarchy of the XML role policy at Edinburgh, any privileges that the external role holds will be inherited by the appropriate roles as deemed suitable by the local

Edinburgh SoA, e.g. an *externalStudent* may have less privilege than an *EdinburghStudent* role which already exists within the Edinburgh PMI. This hierarchical management of roles allows distinct levels of trust to be implemented based on a user's function and location within the VO without surrendering local policy integrity. Thus for example, Glasgow students are able to access Edinburgh Grid compute resources but not allowed to print on local printers.

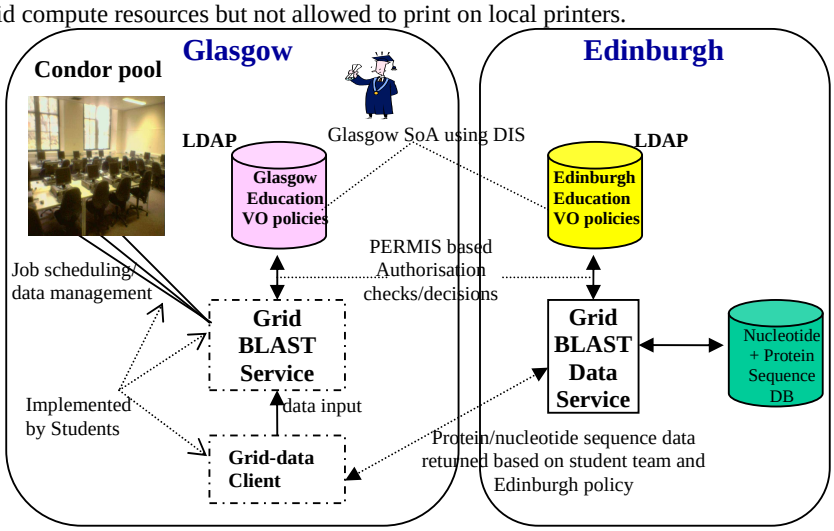

Figure 2: Grid Computing Assignment Utilizing Dynamic AC Creation and Authorization

4. CONCLUSIONS

The dynamic establishment and subsequent management of VOs represents a significant security challenge to the Grid community (if it is done correctly!), but a challenge that needs to be overcome in order for Grid technology to be taken up by the more security focused communities such as the medical domain, or industry more generally. The dynamic delegation of authority infrastructure supported within the DyVOSE project offers one possibility through which advanced authorization infrastructures can be linked dynamically. Through delegated creation of VO specific roles and attributes, VOs can be established in a dynamic manner without compromising the overall security. At the time of writing the students at Glasgow are in the final phases of their implementation work with the successful return of sequence data already completed based upon dynamically created security attributes – thus proving the proof of concept in using DIS for dynamic VO establishment and fine grained authorization.

To simplify the overall process in access to and usage of VO Grid resources, the Shibboleth technology offers direct benefits for single sign-on, but currently requires a more static view of the security attributes that are available. Through the DIS service, an SP may, subject to its having the appropriate privilege at IdPs be allowed to create attributes for those IdPs which will subsequently be needed for access to the resource. This model significantly changes the dynamics through which future VOs may be composed. Truly dynamic security oriented VOs where service providers not only offer services but the attributes needed for access to these services has hitherto not been addressed by the Grid community. This work is being

explored in a variety of security oriented projects at the National e-Science Centre especially in the e-Health domain.

4.1 Acknowledgments

This work was funded from a grant from Joint Information Systems Committee (JISC) in the UK. The authors would like to thank their collaborators in this project in particular Professor David Chadwick and Dr Sassa Otenko at the University of Kent for their work on production of the DIS service.

5. REFERENCES

1. R. Alfieri,et al, Managing Dynamic User Communities in a Grid of Autonomous Resources, CHEP 2003, La Jolla, San Diego, March, 2003;
2. D.W.Chadwick, A. Otenko, The PERMIS X.509 Role Based Privilege Management Infrastructure, Future Generation Computer Systems, 936 (2002) 1–13, December 2002. Elsevier Science BV.
3. D.W. Chadwick, Delegation Issuing Service, NIST 4th Annual PKI Workshop, Gaithersberg, USA, April 2005.
4. Condor project, www.cs.wisc.edu/condor
5. Dynamic Virtual Organisations for e-Science Education, project www.nesc.ac.uk/hub/projects/dyvose
6. eduPerson Specification, http://www.educause.edu/eduperson/
7. I. Foster, C. Kesselman, S. Tuecke, The Anatomy of the Grid: Enabling Scalable Virtual Organizations, International Journal of Supercomputer Applications, 15(3), 2001.
8. Global Grid Forum, (V. Welch, F. Siebenlist, D. Chadwick, S. Meder, L. Pearlman), Use of SAML for OGSA Authorization, June 2004, https://forge.gridforum.org/projects/ogsa-authz
9. Globus Security Infrastructure (GSI), http://www.globus.org/security/
10. Globus toolkit, http://www.globus.org/toolkit/downloads/4.0.1/
11. R. Housley, T. Polk, Planning for PKI: Best Practices Guide for Deploying Public Key Infrastructures, Wiley Computer Publishing, 2001.
12. J. Jokl, J. Basney and M. Humphrey, Experiences using Bridge CAs for Grids, Proceedings of UK Workshop on Grid Security Practice - Oxford, July 2004.
13. OpenSSL, www.openssl.org
14. L Pearlman, et al., A Community Authorisation Service for Group Collaboration, in Proceedings of the IEEE 3rd International Workshop on Policies for Distributed Systems and Networks. 2002.
15. W. T. Polk and N. E. Hastings, Bridge Certification Authorities: Connecting B2B Public Key Infrastructures, http://csrc.nist.gov/pki/documents/B2B-article.doc
16. A. Robiette, T. Morrow, Blueprint for a JISC Production Federation, JISC Development Group, Version 1.1: issued 27 May 2005, http://www.jisc.ac.uk/index.cfm?name=middleware_documents
17. OASIS, Assertions and Protocol for the OASIS Security Assertion Markup Language (SAML) v1.1, 2 September 2003, http://www.oasis-open.org/committees/security/
18. Shibboleth Architecture Technical Overview, http://shibboleth.internet2.edu/docs/draft-mace-shibboleth-tech-overview-latest.pdf
19. Shibboleth Architecture Protocols and Profiles, http://shibboleth.internet2.edu/docs/draft-mace-shibboleth-arch-protocols-latest.pdf
20. R.O. Sinnott, D.W. Chadwick, Experiences of Using the GGF SAML AuthZ Interface, Proceedings of UK e-Science All Hands Meeting, September 2004, Nottingham, England.
21. A.J. Stell, R.O. Sinnott, J. Watt, Comparison of Advanced Authorisation Infrastructures for Grid Computing, Proceedings of International Conference on High Performance Computing Systems and Applications, May 2005, Guelph, Canada.
22. A.J. Stell, Grid Security: An Evaluation of Authorisation Infrastructures for Grid Computing, MSc Dissertation, University of Glasgow, 2004.
23. UK e-Science Certification Authority, www.grid-support.ac.uk/ca
24. J. Watt, R.O. Sinnott, A.J. Stell, Dynamic Privilege Management Infrastructures Utilising Secure Attribute Exchange, Proceedings of UK e-Science All Hands Meeting, Sept. 2005, Nott, England.
25. J. Watt, R.O. Sinnott, O. Ajayi, J. Jiang, J. Koetsier, A Shibboleth-Protected Privilege Management Infrastructure for e-Science Education, to appear in 6th IEEE International Symposium on Cluster Computing and the Grid, CCGrid2006, May 2006, Singapore.
26. ITU-T Recommendation X.509 (2001) | ISO/IEC 9594-8: 2001, Information technology – Open Systems Interconnection – Public-Key and Attribute Certificate Frameworks.