University of Glasgow | Experimental Particle Physics

GLAS-PPE/2008-34
26th November 2008

Department of Physics and Astronomy
Experimental Particle Physics Group
Kelvin Building, University of Glasgow
Glasgow, G12 8QQ, Scotland
Telephone: +44 (0)141 330 2000 Fax: +44 (0)141 330 5881

Advanced Security for Virtual Organizations: The Pros and Cons of Centralized vs Decentralized Security Models

R.O. Sinnott[1], D.W.Chadwick[2], T. Doherty[1], D. Martin[1], A. Stell[1], G. Stewart[1], L. Su[2], J. Watt[1]

[1]*National e-Science Centre, University of Glasgow*
[2]*Information Systems Security Group, University of Kent*
Email: r.sinnott@nesc.gla.ac.uk or t.doherty@physics.gla.ac.uk

Abstract

Grids allow for collaborative e-Research to be undertaken, often across institutional and national boundaries. Typically this is through the establishment of virtual organizations (VOs) where policies on access and usage of resources across partner sites are defined and subsequently enforced. For many VOs, these agreements have been lightweight and erred on the side of flexibility with minimal constraints on the kinds of jobs a user is allowed to run or the amount of resources that can be consumed. For many new domains such as e-Health, such flexibility is simply not tenable. Instead, precise definitions of what jobs can be run, and what data can be accessed by who need to be defined and enforced by sites. The role based access control model (RBAC) provides a well researched paradigm for controlling access to large scale dynamic VOs. However, the standard RBAC model assumes a single domain with centralised role management. When RBAC is applied to VOs, it does not specify how or where roles should be defined or made known to the distributed resource sites (who are always deemed to be autonomous to make access control decisions). Two main possibilities exist based on either a centralized or decentralized approach to VO role management. We present the advantages and disadvantages of the centralized and decentralized role models and describe how we have implemented them in a range of security focused e-Research domains at the National e-Science Centre (NeSC) at the University of Glasgow

Advanced Security for Virtual Organizations: The Pros and Cons of Centralized vs Decentralized Security Models

R.O. Sinnott[1], D.W.Chadwick[2], T. Doherty[1], D. Martin[1], A. Stell[1], G. Stewart[1], L. Su[2], J. Watt[1]

[1]*National e-Science Centre, University of Glasgow*
[2]*Information Systems Security Group, University of Kent*
r.sinnott@nesc.gla.ac.uk

Abstract

Grids allow for collaborative e-Research to be undertaken, often across institutional and national boundaries. Typically this is through the establishment of virtual organizations (VOs) where policies on access and usage of resources across partner sites are defined and subsequently enforced. For many VOs, these agreements have been lightweight and erred on the side of flexibility with minimal constraints on the kinds of jobs a user is allowed to run or the amount of resources that can be consumed. For many new domains such as e-Health, such flexibility is simply not tenable. Instead, precise definitions of what jobs can be run, and what data can be accessed by who need to be defined and enforced by sites. The role based access control model (RBAC) provides a well researched paradigm for controlling access to large scale dynamic VOs. However, the standard RBAC model assumes a single domain with centralised role management. When RBAC is applied to VOs, it does not specify how or where roles should be defined or made known to the distributed resource sites (who are always deemed to be autonomous to make access control decisions). Two main possibilities exist based on either a centralized or decentralized approach to VO role management. We present the advantages and disadvantages of the centralized and decentralized role models and describe how we have implemented them in a range of security focused e-Research domains at the National e-Science Centre (NeSC) at the University of Glasgow.

1. Introduction

Grids and the Grid Computing paradigm provide the technological infrastructure to facilitate e-Science and e-Research. Numerous national and international collaborations have successfully shown how Grid technologies can support a wide range of research including amongst others: seamless access to a range of computational resources [1]; linkage of a wide range of data resources [2]; exploitation of shared instruments such as astronomical telescopes or specialized resources such as visualization servers [3]. Indeed there are few application domains that have not been exposed to and benefited from the exploitation of Grid technologies.

Given this, it would be expected that a body of knowledge exists on how best to apply Grid technologies to support e-Research. It is still the case however that a variety of choices exist in both the interpretation of the Grid computing paradigm, and on the technologies that are used to realize Grid based e-Research infrastructures. Historically, much of the focus and effort of Grid computing was based upon addressing access to and usage of large scale high performance computing (HPC) resources such as cluster computers. These access models are typified by their predominantly authentication-only based approaches which support secure access to an account on a cluster where domain specific programs can be compiled and/or executed. These approaches are based upon X.509 based public key infrastructures (PKI) [4] where the public key certificates (PKCs) that bind the identities of users to their public keys are issued by trusted third parties called certification authorities (CAs). Through trusting a CA, sites can validate the identity of the individual in possession of the corresponding private key. This PKI based approach has been adopted in the UK by the National Grid Service (NGS) (www.ngs.ac.uk) with the UK certification authority based at Rutherford Appleton Laboratory (www.grid-support.ac.uk/ca). With this model, end users are expected to acquire and manage their own private/public key pairs. This is an arduous process for many less IT-savvy communities who are put-off by the acquisition and management of X.509 PKCs, e.g. they may be expected to memorize 16-character long private key passwords. The temptation to write down such passwords or share them between users also results in a reduction of overall security. These issues are described in more detail in [5-7].

It is often the case that research domains and resource providers require more information than simply the identity of the individual in order to grant access to use their resources. The same individual can be in multiple collaborative projects each of which is based upon a common shared infrastructure. Knowing in what context a user is requesting access to a particular resource is essential information for a resource provider to decide whether the access request should be granted or not. This information is typically established through the concept of a virtual organization (VO). A VO allows the users, their roles and the resources they can access in a collaborative project to be defined. This information can then be used by individual resource providers to decide upon the validity of access requests, e.g. through satisfaction of their site specific authorization policies.

There are numerous technologies and standards that have been put forward for defining and enforcing authorization policies for access to and usage of Grid resources [8]. Role based access control (RBAC) is one

of the more well established models for describing such policies [26], although other models such as attribute based, process based, and identity based access control also exist and have been implemented [10]. In the RBAC model, project/VO specific roles are assigned to individuals as part of their membership of a particular VO. Possession of a particular role, combined with other context information, such as time of day and amount of resource being requested, can then be used by a resource gatekeeper to decide whether an access request is allowed or not. RBAC has numerous advantages over existing authentication only based models. One of the key advantages is that whilst individuals in a VO may come and go, the role itself is unlikely to change as much. Consequently RBAC based approaches are considered more scalable and manageable. The key advantage of RBAC-based security models compared to other approaches is that privileges and access is determined by roles and memberships a user holds and not merely on identity. However, the ANSI RBAC model [26] assumes a single domain with centralized role management so that conflicting roles cannot be issued to users and all systems know which roles a user is a member of. These assumptions do not necessarily hold true in VOs.

Many mainstream Grid infrastructures such as the UK e-Science NGS have implemented access control based primarily on X.509 PKCs. In this model, users specifically request access to individual NGS resources by quoting their Distinguished Name (DN) which is embedded in their X.509 PKC. Their DN is registered in a resource maintained grid mapfile which associates their DN with a local account on that resource. If a DN does not have an associated local account then the local gatekeeper will decide that the user does not have privileges to run the job. RBAC models instead allow possession of particular roles to determine access control decisions.

It is not realistic to mandate that all resource providers should adopt precisely the same security policies, nor how end users with different roles should access and use grid resources. Each site can/will have their own mechanisms for dealing with site specific access requests. Indeed the common philosophy underlying the Grid is that all resource providers are expected to be autonomous, i.e. they may allow/deny access requests at their own discretion. Nevertheless, a crucial consideration in establishing a VO is whether a common understanding of the various roles and their associated privileges needs to be established throughout the entire VO or not. There are two primary models for defining roles specific to a VO: the centralized and decentralized models. The focus of this paper is to explore these two models and identify their advantages and disadvantages through application in security focused Grid projects. Key to this is to ensure that the models are trialled in realistic, large scale heterogeneous Grid environments. To this end we have been working with mainstream Grid technologies such as Globus toolkit version 4 (www.globus.org/toolkit) and the Open Middleware Infrastructure Institute (OMII-UK) software (www.omii.ac.uk). Seamless support of the two VO role models with Grid middleware is essential, and is something we pay special attention to.

2. Pros and Cons of Centralized vs Decentralized VO Role Models

The centralized model more nearly matches the standard RBAC model, and requires all sites to agree upon the roles and privileges that are to be used throughout a particular VO. In this model, all sites agree in advance on the definition and names of the roles that are applicable to their particular VO, and the privileges that will be assigned to them. A single VO administrator is then appointed who will typically assign these roles to individuals on a case by case basis when users ask to be granted particular roles or permissions in the VO. The VO administrator may appoint other administrators to help him in this task, but all administrators are conceptually equal, in that each can over-ride the decisions made by the others. This model of VO role administration has been implemented through technologies such as the Virtual Organization Membership Service (VOMS) [11]. VOMS has gained widespread acceptance due to the simple model for defining the roles specific to a particular VO and how they can be used/enforced. Sites themselves are responsible for configuring their resources to use these roles. With VOMS, this is implemented with tools such as the Local Centre Authorization Service (LCAS) and the Local Credential Mapping Service (LCMAPS) [12] which map the user role information into group identities (*gid*), user identities (*uid*) and associated local pool accounts established on the local cluster for that particular VO. Refinements can be made to this model in order to allow more local control over the use of resources, e.g. applying file store limits to a particular VO. We note that this local enforcement is not explicitly defined within the VO policy (given by the definition of the roles in the VOMS server). Rather, this is left up to local administrators to decide how the particular roles and privileges associated with that VO should be interpreted when accessing the resource.

The decentralized VO role model is more aligned with the original dynamic collaborative nature of the Grid as first put forward [1]. In this model there is no central VO administrator. Instead, each resource site has its own local administrator who is 100% responsible for determining which VO members can access the local VO resources. Each site administrator determines the roles and the associated privileges that are required to access and use the local resources. However, it would not be realistic or scalable to expect each site administrator to assign the various roles to all the users in the VO. Consequently, they can decide which other administrators (at this and other VO sites) are trusted to assign which roles to which VO users. In this way they may each delegate to each other the responsibility of user-role assignments throughout the VO. However the

assignment of privileges to roles is always under their direct control in their local policy. In the decentralised approach, no centralized agreement takes place. Rather, dynamic peer to peer collaborations are supported by site administrators delegating the necessary privileges to the site administrators (and users) of collaborating organizations. Underpinning this model are bilateral trust agreements between sites. This is the model on which the PERMIS authorisation system [9] has been built. One can see that the centralised model is a subset of the distributed model, in which all sites bilaterally trust a single VO administrator.

Both the centralized and decentralized approaches have their advantages and disadvantages. We summarize these below and give an outline of when each approach is beneficial. In both cases we consider RBAC as the mechanism for access control but the principles underlying the centralised or decentralised models are broadly applicable to some other approaches as well, e.g. attribute based access control.

2.1. Pros of Centralized VO Role Model

The centralized VO role model, such as that based on VOMS, has several advantages. Firstly it is widely accepted across the Grid community e.g. VOMS has been accepted by many large scale mainstream HPC-oriented Grid communities. Consequently good tool support exists for the central management of these roles, such as VOMRS [24] which allows multiple managers to assign roles to members of the VO, and for end users to see which roles they have been allocated. Furthermore, tools such as *voms_proxy_init* exist for embedding these roles into proxy certificates and for pushing them to the resource sites. Other complementary tools exist for extracting the roles from the proxy certificates at the resource site.

The centralized VO model is ideally suited when large scale, primarily static VOs are needed. Here static implies that the roles and end users with those roles do not change rapidly across the VO. The interpretation and mapping of those roles to local resources may well change more frequently however. Thus through technologies such as LCMAPS/LCAS a local system administrator is able to decide which local pooled accounts and file storage is made available to members of that VO in a dynamic manner. If a user's privileges are to be revoked, then the VO administrator can simply remove the roles assigned to this user in the VOMS server, with the consequence that the user's roles are no longer recognized across the whole VO.

Given that the VO roles are agreed by all sites up front when establishing the VO, the centralized model is simpler to define and agree upon. This model does not depend on the aggregation of numerous bilateral agreements between VO partners where roles and associated trust levels are defined. Rather roles are defined globally across the VO, based upon a VO-wide collaborative agreement. The assignment of these roles to individuals is then made by a designated VO-manager – typically the VOMS administrator (although the manager role can be shared by several people). This super-role is responsible for deciding which users can be assigned which roles across the VO. The knowledge of all possible users involved in the VO and their roles implies a lack of scalability with this model. However, we note that moderately large VO infrastructures have been established adopting this model. For example, the VO for the ATLAS project (http://www.usatlas.bnl.gov/) has over 1500 members with a variety of different roles.

The centralized VO role model, or more precisely agreement on a core set of roles, is also aligned with the principle behind the definition of the *eduPerson* attribute set (www.educause.edu/eduperson/) for use with technologies such as the Internet2 Shibboleth (shibboleth.internet2.edu). Through widespread definition and agreement of the roles to be used across a federation, these may then be delivered and used in a variety of ways.

The centralized role model is also well aligned with HPC-oriented models of Grid usage. Thus, mapping of VOMS user roles to local pooled accounts on large scale clusters is their typical *modus operandi*. Restricting access to particular data sets is typically not required. The centralized model is also quite flexible in that the low level detailed policy definition is left open to interpretation by individual sites. Thus a site may decide whether given roles will be recognized or not, and if so how, e.g. what pooled accounts they should be mapped to. Given that LCMAPS/LCAS have been targeted at the pooled account level, this means that this model of security supports communities who wish to develop their own programs and run them across clusters (albeit in given VO-specific accounts). Thus this model supports *tinkerers* and HPC-oriented communities who do not simply require access to known services.

2.2 Cons of Centralized VO Role Model

The centralised VO role model based on VOMS also has its disadvantages. Having a single VOMS server is a single point of failure. Should this resource become unavailable for whatever reason, then no resources across the VO will be accessible.

Having a centralised VO administration model also has potential drawbacks. For larger scale VOs, it is unlikely that a single administrator will have the detailed knowledge to decide whether a given remote end user should have a particular role or not. For smaller scale VOs this may not be a problem but as VOs scale both in terms of their number of users and the frequency at which privileges are assigned/revoked or new privileges added, this model becomes more difficult to scale. It is for this reason that tools such as VOMRS have added support for multiple people to perform the role of VOMS administrator. Note however that each role occupant is a full VOMS administrator and can therefore undo or redo the role assignments of other administrators. Consequently, conflicts between administrators have to be solved outside the model.

Whilst it is relatively straightforward for a VOMS-administrator to add a new role to an existing VO, the roll-out and interpretation of this by all the resource sites e.g. through LCMAPS/LCAS mapping this new role to an appropriate local account, may still cause scalability issues.

2.3 Pros of De-centralized VO Role Model

The decentralised model of VO's has several advantages. Firstly, it allows for more dynamic collaborations to occur. Thus rather than all sites having to agree on VO-wide roles and develop associated policies, the decentralised model allows a resource administrator to directly provide end users and trusted end user administrators with the privileges they need to enable access to his resource. New VO roles do not need to be created and assigned to end users, their existing roles can be used.

Supporting this model only requires trust to exist between pairs of collaborating administrators and this is probably much easier to establish than trust between every administrator and a single external centralised VO administrator. It may be simply realised through delegation of authority, whereby a resource administrator delegates to a remote (trusted) administrator the privilege to issue to a subset of VO users a subset of the roles needed to access their resources. The Delegation Issuing Service (DIS) from the Dynamic Virtual Organizations for e-Science Education (DyVOSE) project [25] (www.nesc.ac.uk/hub/projects/dyvose) provides one implementation of such a delegation of authority model.

The decentralised model is more scalable than the centralised model since there is no limit to the number of remote trusted administrators that can be delegated the task of assigning roles to users, and since each has a limited scope of operation, there is no fear of any administrator undoing the work of another one.

The decentralised model is also more reliable in associating roles with users, since the assignment of roles can be done by remote administrators who are based at the same sites as the users, and who therefore can be expected to know the users better. This is similar to the motivation behind the registration authority (RA) concept in PKIs. Thus rather than adopting a single central VO-administrator to assign all roles, each site involved in the collaboration may appoint their own distributed set of administrators to assign different roles to different groups of users at different sites. Each administrator privilege can be independently granted and denied by each resource administrator, thereby providing a fine level of granularity for user-role assignments. Each site will consequently have its own set of local VO-administrators who are responsible for assigning to local VO-members the various roles that are needed to access resources throughout the VO, i.e. it is the local VO-administrators who are authorised by the remote resource providers to assign the VO-specific roles to their local users. Each site VO administrator is therefore considered to be an Attribute Authority (AA), and multiple AAs may therefore exist both for each VO role and at each VO site.

The model also has the advantage of being more tolerant to partial failures. Thus should any single AA fail, the VO itself will persist and some/most end users will still be able to access some or all of the VO resources. Of course, if the AA that fails is the sole provider of the roles for a specific resource then VO fault tolerance cannot be guaranteed.

The model also has the advantage that new VO roles do not always need to be created. If an organisation joins a VO, and an existing group of employees within the organisation are all to become members of the VO, then the role that confers the existing group membership may be used by the VO to grant access to this group to VO resources. This simply requires a VO resource administrator to recognise (or add) this role into its policy and assign the appropriate permissions to this already-existing role. This is in fact how credit card authorisation works today. When a new retail outlet (resource) comes on line, it simply trusts the credit card roles that already exist.

The decentralised VO model is much more flexible. New resources and roles can be provided on the fly, in an incremental fashion between collaborating partners. Similarly sub-groups of users can be excluded on the fly by removing trust from their administrator (AA) who assigned their roles, without removing the role from either the VO or from other users who have had the same role assigned by other (still trusted) AAs. Sites may make their own local decisions, based on whatever local information is to hand, whether to define new roles, use existing roles, change or revoke existing roles, trust new administrators, or give greater or less trust to existing administrators etc.

Decentralised models based upon technologies such as PERMIS allow for finer grained access control to be supported – in comparison to the VOMS based approach of mapping a user role to a particular gid/uid and pooled account. Instead PERMIS allows for policies based upon the combination of *roles, targets* and *actions* to be defined. A typical scenario here is "can this user with this role access this service and invoke this method". Through definition of such site specific policies finer grained access control can be defined. Thus end users will not typically be *tinkerers* as described in section 2.1 but service users.

Tool support now exists to support the decentralised model of VO roles. The PERMIS toolset in particular has extensions to allow the secure creation and delegation of roles which directly map onto the decentralised VO role model.

2.4 Cons of De-centralized VO Role Model

The decentralised model of VO roles is not without its disadvantages. One of the major problems with this model is that the roles associated with the VO are potentially scattered across numerous locations, being provided as they are by numerous AAs. Where should users go to in order to obtain the necessary roles that are

needed to grant them access? This problem is less severe with the centralised model, which only has a central server, although a user who is a member of several centralised VOs still has some choices to make. One solution is for resource providers to support the pull model of role collection since it removes the burden from the end users. Since the resource administrators determine which AAs they trust to issue which roles to whom, they are able to configure AA meta-data into their resource gatekeepers that provide instructions where to go to, in order to pull the various user roles that are needed. But a consequence of this is that maximum rather than least privileges results. Achieving least privileges with user push of roles will require either all users to know where all of their roles are located, or the AAs to distribute the signed roles to their users, as is done today with public key certificates.

Another issue with decentralised models is how do the VO-specific roles get defined and made known to the sites where they need to be assigned to users? They could be defined collaboratively by resource provider and user sites, or VOs could centrally define all the roles that will apply throughout the VO, but the model does not require this. One approach to achieving this is the Delegation Issuing Service (DIS) developed within the DyVOSE project. The DIS is a web service that allows role assignments and delegation to be performed from anywhere by anyone who can successfully authenticate to it (via PKI) and who has the necessary permissions. An Apache server front end to DIS allows users without PKI credentials to authenticate to Apache which then acts as a proxy for the user. The DIS allows role management to be delegated to other administrators in the privilege management infrastructure. The application of the DIS to decentralised role management was validated in the e-learning domain, but it can be generalised to other application domains. More information on the DIS is given in [13].

One final non-technical disadvantage with the decentralised VO role model is that it represents a new paradigm for the mainstream Grid community. As such, its uptake and exploitation has not been as great as with VOMS. However, distributed role management is the reality of the world today, and given the supporting tools and move towards Shibboleth based access to Grid resources through projects such as GridShib (gridshib.globus.org), ShibGrid [14] and Shebangs [15] and GLASS (www.nesc.ac.uk/hub/projects/glass), the decentralised model is becoming increasingly relevant.

The ideal scenario might be to combine the benefits of the centralised VO model, ala VOMS (which has wide adoption) with technologies that support the decentralised model in order to provide a hybrid approach where either or both models can co-exist together, thereby providing a finer grained, scalable and more manageable access control infrastructure. Within the JISC funded VPMan project (http://sec.cs.kent.ac.uk/vpman/) we are exploring how this can be achieved – specifically through the combination of VOMS and PERMIS.

3. Implementation of a Hybrid Centralised-Decentralised VO Role Model

The initial focus of the VPMan project is to combine VOMS and PERMIS to provide a centralised VO role management system with the fine grained access controls of PERMIS. We recognised that integration with existing and widely adopted Grid middleware is essential in order to painlessly migrate users to more flexible solutions. The secondary focus will be to introduce distributed role management so that users will be able to successfully combine their roles which have been issued by different distributed AAs. This will be a generalisation of the specific solution currently being implemented in the Grid-Shib project. In our implementation efforts thus far, we have focused predominantly upon implementation of scenarios within the Globus toolkit version 4 (GT4) and with OMII-UK.

3.1 GT4 Implementation

The Globus technologies have had widespread adoption by the Grid and e-Science communities. The latest version of Globus is GT4. The authorization framework associated with GT4 provides capabilities to plugin a series of interceptors to process each request when it is received, i.e. before it reaches the protected application. Two types of interceptors are of interest from an authorization perspective: Policy Information Points (PIPs) and Policy Decisions Points (PDPs). The main task of a PIP is to prepare an appropriate component of the request context ready for it to be passed to the PDP for an access control decision. Typically there will be a PIP to prepare each of: the subject's attributes, the action's attributes, the resource's attributes and the environmental attributes. The relationship between PIPs, PDPs and the Policy Enforcement Point (PEP) with GT4, PERMIS and VOMS is shown in Figure 1.

VOMS is integrated with the Globus Toolkit so that the user's roles encoded as X.509 attribute certificates (ACs) can be passed around embedded in X.509 proxy certificates. A GT4 VOMS PIP allows GT4 to access and process the VOMS ACs (the Subject PIP in Figure 1). The VOMS PIP extracts the VOMS ACs from the proxy certificate, parses and stores the roles in the GT runtime so that they may subsequently be used by PDPs for making authorisation decisions.

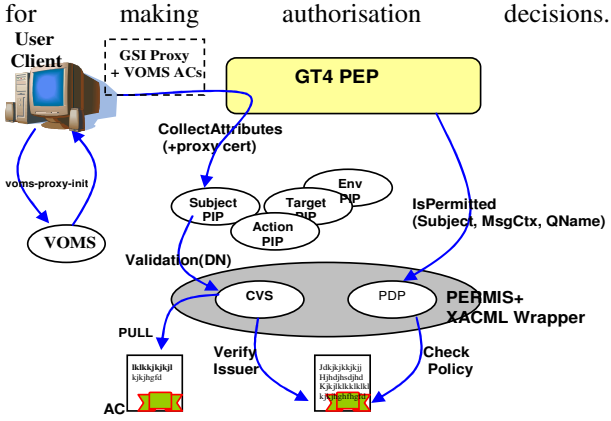

Figure 1: GT4 VOMS-PERMIS Integration

The PERMIS Credential Validation Service (CVS) intercepts these roles and uses its policy to decide if they were correctly assigned by a trusted AA. In the current policy there is only one trusted AA, the centralised VOMS server, but in future policies there will be a distributed set of trusted AAs. The PERMIS CVS has the ability to pull additional roles from distributed AAs and merge them with the pushed ones, and this will be used to support the distributed role model. VOMS roles may then be picked up from a VOMS SAML service given the DN of the user. The valid set of user roles is passed back to the GT runtime for passing to the PDP (or other PIPs, depending upon the GT4 configuration).

Current PIP implementations usually package the various subject, action, resource and environmental attributes in a standard XACML request context format [16] so that they can be passed to any XACML conformant PDP. The PERMIS PDP also has an XACML wrapper interface, so that it can be swapped for an XACML PDP when needed.

To actually secure a GT4 service, it should be configured so that the required PIPs as well as a PDP must be called before access is granted. These PIPs will create the various components of an XACML request context and once all required information is collected, the PDP is passed a completed XACML request context. A protected GT4 service is configured with a security configuration and a service configuration. The former indicates the authorisation and authentication methods. In the authorisation method description, the PIPs and PDP are specified in the format of <identifier>:<java module> where *identifier* specifies a certain scope and *java module* is the full name of the java module which implements a PIP or PDP. The identifier for a PIP/PDP is used to differentiate between module instances and the parameters that need to be passed to each instance. Other services may use the same modules but with different configurations by using different identifiers. We note that the system has been designed to be extensible so that other PIPs or PDPs may be added to the authorisation chain.

3.2 OMII-UK Implementation

OMII-UK was created to establish and maintain Grid middleware for the UK e-Science community. The OMII software stack incorporates a rich set of software for service development, discovery and management, for workflows based on the Taverna workflow environment (www.mygrid.org.uk) and for management of e-Science data sets through technologies such as OGSA-DAI (www.ogsadai.org.uk). A variety of newly commissioned projects have also been funded identifying particular needs for the wider research community, e.g. for visualisation (www.nesc.ac.uk/hub/projects/omii-rave) or finer grained security via Shibboleth (www.nesc.ac.uk/hub/projects/omii-sp).

The currently supported security model within the OMII-software stack is primarily based upon web service security models. These are used to provide secure access to a variety of services including GridSAM (gridsam.sourceforge.net) which is itself targeted at job submission and monitoring across a range of computational resources. The OMII-AuthZ project has provided an implementation that supports the OGF AuthZ SAML callout API [17]. Details of how this API was both linked to Grid middleware and exploited more generally are described in [8].

The system architecture depicting how VOMS and PERMIS are being integrated within the OMII-UK technologies is depicted in Figure 2.

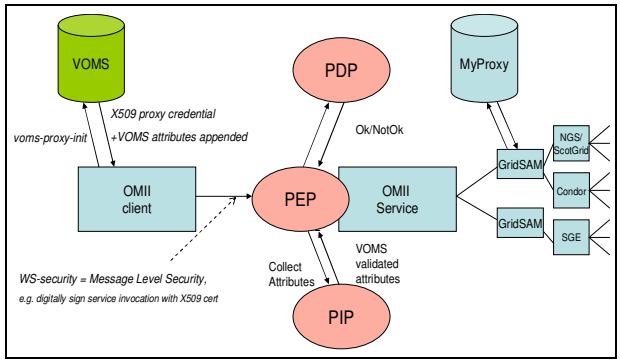

Figure 2: OMII-UK VOMS-PERMIS Integration

In this infrastructure, the typical scenario is where an end user creates an X509 proxy certificate with a VOMS attribute certificate embedded, either through *voms-proxy-init* or via the Acacia software [18]. Upon attempting to invoke a PERMIS protected service, the VOMS attribute certificates are extracted from the X.509 proxy certificate which is transferred as part of the Job Submission and Description Language (JSDL) [19] message and validated using the PERMIS CVS and PIP through similar mechanisms described previously in the GT4 scenario. These VOMS attributes are then used by the PDP to decide which end resources the job has permission to be submitted to.

4. Case Studies

To demonstrate how we have incorporated the advantages of the centralised VO role model with technologies primarily established for decentralised VO role management, we focus on two case studies: the MRC funded pilot project *Virtual Organizations for Trials and Epidemiological Studies* (VOTES - www.nesc.ac.uk/hub/projects/votes) and the EPSRC funded pilot project *Meeting the Design Challenges of nanoCMOS Electronics* (nanoCMOS – www.nanocmos.ac.uk).

4.1 Clinical Trials and Epidemiological Domain

Clinical trials and clinical systems more generally place many demands upon security infrastructures to support the various activities involved. In particular, the typical processes involved in a clinical trial will comprise recruitment, collection of data specific to the trial and

overall management of the trial itself, e.g. to ensure that it is undertaken according to ethical concerns.

Fine grained security is essential in this context to ensure that the right data is made available to the right people for the right purpose. A key aspect of the work is that VOTES is not concerned with developing a single Grid infrastructure for a specific clinical trial or study, but with developing a Grid based framework through which a multitude of clinical trials can be supported. There is a multitude of clinical resources already available containing clinical information across Scotland including: GPASS - used by 85% of GPs; Scottish Care Information store - used by most health trusts across Scotland; and Scottish Morbidity Records - containing aggregated clinical data over 25 years including cancer registrations, disease registries and death data sets to name just some. The architecture used within VOTES is described in more detail in [20-22].

The typical security mechanism in VOTES involves a call-out to a third party database containing the various access rights that the different roles within the infrastructure have. Conceptually, the access matrix represents the privileges assigned for the entire Clinical Virtual Organisation (CVO), but is in essence an aggregation of all the local resource access policies. This relates to the centralized model of security, where one policy describes all nodes. In practice however, this is a difficult model to implement, partly because of the heterogeneity of the different data resources, and partly because of a lack of trust between partners in adopting a new security system. As such, a requirement of this project is to be able to combine security policies of local resources, with an overarching policy that requires a pluggable interface, rather than subscribing to a set global schema.

Versions of this framework utilize GT4, OGSA-DAI, GridSphere and PERMIS. To exploit VOMS and PERMIS, a new trial was established focused upon support of a diabetes clinical study. This trial identified two key roles: *VOTESdiabetes-doctor* and *VOTESdiabetes-nurse*. These roles were recorded in a *VOTESdiebetes* VO maintained in a VOMS server at NeSC. The data sources used in this study included SCIstore; the Community Health Index database and a Consent database for patients who have agreed to partake in the study. The roles provide access to fixed pre-agreed queries (stored procedures) which must strictly adhere to the trial protocol. To show how VOMS and PERMIS could be combined five separate stored procedures were implemented providing access to either demographic but non-clinical data (for *nurses*) or for more clinically oriented data (for *doctors)*. We note that these roles are specific to the *VOTESdiabetes* trial only. To understand how VOMS/PERMIS work together, we consider the typical sequence of steps:

1. A user creates an X.509 proxy certificate with appropriate VOMS attributes embedded, e.g. through voms-proxy-init as below:

```
voms-proxy-init –voms \
votesdiabetes:/votesdiabetes/Role=nurse-cert
```

This step can also be achieved using the Acacia tool.

2. The user tries to invoke the service that gives access to the protected stored procedure;
3. This invocation is intercepted by the PEP which passes the user information including the proxy certificate and appended attributes to the VOMS PIP;
4. The VOMS PIP extracts and validates the credentials and passes back a VOMS Fully Qualified Attribute Name (FQAN) with the subject attributes;
5. The PEP calls the PERMIS PDP pushing the request information and credentials;
6. The PERMIS PDP will then, according to the policy, decide if this user with the presented attributes is authorized to access the service;
7. If they are authorised the stored procedure is invoked, the federated query run and the resultant data sets joined on the CHI number and returned to the end user.

This system has shown how VOMS and PERMIS can be combined to provide secure access to federated clinical data, however many domains require secure fine grained access to HPC computational resources.

4.2 nanoCMOS Electronics Domain

For next generation nanoCMOS electronics design, the quantum level effects of devices of ever decreasing dimensions are becoming ever more important and atomistic simulation of devices is essential. The nanoCMOS project is developing an infrastructure through which device level designs and simulations can be linked through to higher level circuit and system simulations, to predict the overall behaviour of nanoCMOS systems. However, this domain demands infrastructures that support protection of intellectual property, be it for designs of transistors, data sets, simulation codes or circuit/systems designs, hence fine grained security is essential.

For the development of the Grid infrastructure the project has aligned itself with the OMII-UK technologies. The early phase of the work focused upon developing a family of OMII-UK services which support the device modelling and compact model generation phases of electronics design. These services were developed to exploit the OMII-UK GridSAM job submission system.

To support VOMS integration in the nanoCMOS domain, a *nanoCMOS* VO was established in a VOMS server at NeSC. In this, several key roles were established: *deviceModeller* and *circuitSimulator* roles. These roles were used within vanilla VOMS scenarios to map end users within the nanoCMOS domain to appropriate pooled accounts and gids/uids for the nanoCMOS project. This used vanilla VOMS scenarios with the ScotGrid (www.scotgrid.ac.uk) resource.

In addition, work is on-going in providing authorisation capabilities to GridSAM itself. The aim of GridSAM is to provide a web service for submitting and monitoring jobs managed by a variety of Distributed Resource Managers (DRM). This web service interface allows jobs to be submitted from a client in a JSDL document and supports their status retrieval as a

chronological list of events detailing the state of the job. GridSAM translates the submission instruction into a set of resource-specific actions: file staging, launching and monitoring using DRM connectors for each stage. A variety of resource specific DRM connectors are available including connectors for Condor, Sun Grid Engine and Globus. The work is currently focused on supporting the Globus DRM connector for the *GRAMSubmissionStage* part of the DRM connector sequence. Here authorisation is decided before the JSDL document is submitted to the GridSAM instance and converted to a Globus specific Resource Specification Language document and submitted to a GRAM manager. This is achieved through extraction of the VOMS attributes from the GridSAM invocation (themselves embedded in the JSDL document) and using these to authorise access to specific connectors.

However one issue we have identified in the realisation of this model of security is the need for both service level and resource level security. That is, the authorisation step at the GridSAM::DRM connector level will ensure that only authorised individuals can submit to the remote resources accessible via those DRM connectors. However, it is ultimately the end resource itself, e.g. the NGS nodes, which need to make authorisation decisions. Thus with GridSAM-only authorisation, jobs at the back end appear as "normal" jobs when submitted to the NGS resources, i.e. basic Globus jobs submitted using RSL syntax. To address this we are looking at transferring VOMS information with the job itself. In this case, the authorisation will be made both by the protected GridSAM service and the remote resource provider itself in mapping the request via LCMAPS/LCAS to the appropriate pooled accounts.

5. Conclusions

The models and implementations presented in this paper present alternatives ways that VO security models for Grids can be built and their advantages and disadvantages. Both the centralised and distributed models have their advantages and disadvantages which we have enumerated. We have also shown how centralised VO models can be harmonised with decentralised approaches to gain the best of both worlds. This has been shown to be viable in leading Grid middleware in large scale security-oriented Grid projects. We note that scenarios supporting federated role management/access control have been described in detail in our work within the e-Learning domain [6] as part of the JISC funded DyVOSE project.

One aspect of our work that we are acutely aware of is the need to shield end users from the complexity of the underlying Grid technologies, and also to educate system administrators on how best to establish and maintain secure VOs. Our research is still on-going and much more remains to be done. Different kinds of VOMS-SAML push/pull models for attributes needed for PERMIS based security, and alignment with Shibboleth based access ideas are under exploration where trusted identity providers, i.e. from sites within the VO, are used for authentication and to provide VOMS attributes used for authorisation by service providers. The scoping of trusted identity providers and the attributes they provide has already been demonstrated within the OMII SPAM-GP project at NeSC Glasgow [23].

5.1 Acknowledgements

The VPman, VOTES, nanoCMOS projects are funded by JISC, MRC and EPSRC respectively.

6. References

[1]. I. Foster, C. Kesselman, S. Tuecke, *The Anatomy of the Grid: Enabling Scalable Virtual Organizations*, International Journal of Supercomputer Applications, 15(3), 2001.

[2] R.O. Sinnott, et al, *Grid Services Supporting the Usage of Secure Federated, Distributed Biomedical Data*, UK e-Science All Hands Meeting, September 2004, Nottingham, England.

[3] Astrogrid project, www.astrogrid.org

[4] R. Housley, T. Polk, *Planning for PKI: Best Practices Guide for Deploying Public Key Infrastructures*, Wiley Publishing, 2001.

[5] R.O. Sinnott, et al, *Experiences of Applying Advanced Grid Authorisation Infrastructures*, Proceedings of European Grid Conference (EGC), June 2005, Amsterdam, Holland.

[6] R.O. Sinnott, A.J. Stell, J. Watt, *Experiences in Teaching Grid Computing to Advanced Level Students*, Proceedings of CLAG+Grid Edu Conference, May 2005, Cardiff, Wales.

[7] J. Watt, et al, *Dynamic Privilege Management Infrastructures Utilising Secure Attribute Exchange*, Proceedings of UK e-Science All Hands Meeting, September 2005, Nottingham, England.

[8] 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.

[9] D.W.Chadwick, et al *Role-based Access Control with X.509 Attribute Certificates*, IEEE Internet Computing, March-April 2003.

[10] B. Lang, et al, *Attribute Based Access Control for Grid Computing*, info.mcs.anl.gov/pub/tech_reports/reports/P1367.pdf.

[11] Alfieri R, et al. *VOMS: an authorization system for virtual organizations,* 1st European across Grids conference, Santiago de Compostela.

[12] Local Centre Authorization System, http://hep-project-grid-scg.web.cern.ch/hep-project-grid-scg/lcas-lcmaps.html

[13] R.O. Sinnott, J. Watt, D.W. Chadwick, J. Koetsier, O. Otenko, T.A. Nguyen, *Supporting Decentralized, Security focused Dynamic Virtual Organizations across the Grid*, 2nd IEEE International Conference on e-Science and Grid Computing, Amsterdam, Dec 2006.

[14] D. Spence et al, *ShibGrid: Shibboleth Access for the UK National Grid Service,* Proceedings of UK e-Science AHM, September 2007.

[15]SHEBANGS, www.rcs.manchester.ac.uk/research/shebangs

[16] OASIS *eXtensible Access Control Markup Language (XACML) Version 2.0*, OASIS Standard, 1 Feb 2005.

[17] V. Welch, R. Ananthakrishnan, F. Siebenlist, D. Chadwick, S. Meder, L. Pearlman, *Use of SAML for OGSI Authorization*, GFD.66. March 2006

[18]Acacia,www.gridpp.ac.uk/posters/currentset/metadata_april07.pdf

[19] JSDL, www.gridforum.org/documents/GFD.56.pdf

[20] A. Stell, et al - *Security oriented e-Infrastructures supporting neurological research and clinical trials*, ARES Proceedings, Vienna, Austria, Apr 2007.

[21] A. Stell, et al – *Secure, Reliable and Dynamic Access to Distributed Clinical Data,* LSGrid conference, RIKEN Institute, Yokohama, Japan, Oct 2006.

[22] A. Stell, et al, *Supporting the Clinical Trial Recruitment Process through the Grid*, UK e-Science All Hands conference, Nottingham, UK, Sep 2006.

[23] J. Watt et al, *Federated Authentication and Authorisation for e-Science*, Proceedings of APAC 2007 conference, Perth, Australia, September 2007.

[24] VOMRS, www.uscms.org/SoftwareComputing/Grid/VO

[25] DyVOSE project, www.nesc.ac.uk/hub/projects/dyvose

[26] ANSI Information technology - Role Based Access Control, ANSI INCITS 359-2004.