

Demonstration of Shibboleth in Action across a Range of Security focused Grid Projects

Prof. R. O. Sinnott

National e-Science Centre, University of Glasgow

r.sinnott@nesc.gla.ac.uk

1. Introduction

One of the critical factors to the success of Grid technologies is ease of use. To encourage wider uptake, the access to large scale computational and data resources such as the National Grid Service (NGS) (www.ngs.ac.uk) needs to be made as simple as possible for the end user. Currently, the end user experience of interacting with such resources typically begins with obtaining an UK e-Science X.509 certificate issued by the UK Certification Authority (CA) at Rutherford Appleton Laboratories (RAL) (www.grid-support.ac.uk/ca). This can often be an arduous process, especially for non-IT experienced researchers, requiring them to follow a detailed recipe for obtaining the certificates and converting them into appropriate formats before they are then able to access the resources. Experiences indicate that users are uncomfortable with certificates, their management and overall use to access and use Grid resources. It is also the case that these certificates are primarily used for authentication only whereas many domains require finer grained security models due to the nature of the research, the computational resources, the data sets or for example the fact that licenses are needed to access Grid resources.

The UK academic community is currently in the process of deploying Shibboleth technologies to support local, existing methods of authentication for remote login to resources. Through this model, sites are expected to trust local security infrastructures for example in establishing the identity of users (*authentication*) and their associated privileges (*authorisation*). To support this, the Shibboleth architecture and associated protocols identify several key components that should be supported including Identity Providers (typically their home sites/institutions), Service Providers (for example Grid services or data resources) and optionally Where Are You From (WAYF) services. Through these components, end users will have - ideally - single usernames and passwords at their home sites 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 (authorise) what resources authenticated users are allowed access to.

2. Demonstration

This demonstration will show how the National e-Science Centre at the University of Glasgow has successfully applied Shibboleth across a range of projects to provide fine grained access to a variety of resources in a range of projects. These include:

- MRC funded pilot project: Virtual Organisations for Trials and Epidemiological Studies (VOTES) project;
- JISC funded Dynamic Virtual Organisations for e-Science Education (DyVOSE) project;
- JISC funded Glasgow early adoption of Shibboleth (GLASS) project;
- DTI funded Biomedical Research Informatics Delivered by Grid Enabled Services (BRIDGES) project;
- EPSRC funded pilot project Meeting the Design Challenges of nanoCMOS electronics.

We provide an overview of these projects and give an outline of what each demonstration and associated presentation will entail.

2.1 VOTES demonstration

The VOTES project began in October 2005 and is building a Grid framework through which a multitude of clinical trials and epidemiological studies can be supported. A key part of this is development of fine grained security models which allow for a variety of data access and integration possibilities across a range of clinical data sets. The VOTES work has already implemented proof of concept prototypes which show how Grid middleware such as Globus toolkit version 4, OGSA-DAI and GridSphere can be used to link a variety of clinical data resources across Scotland including:

- Scottish Care Information Store (SCISore) – a batch oriented system used by NHS hospitals to store a range of clinical data sets from pathology records, to admittance records, to biochemistry results;
- General Practitioners Administrative Software System (GPASS) – used by over 85% of GPs across Scotland;
- Scottish Morbidity Records – including a range of health information such as acute events, cancer cases, death data sets including many others;
- Consent database – a critical aspect of conducting a clinical trial is obtaining consent from patients that their data sets may be used for a given purpose;
- Usage of the Community Health Index (CHI) number to link multiple data sets together (this is similar to the NHS number and is unique for each individual in Scotland).

This demonstration will show how Shibboleth can be used to provide a range of data access and integration possibilities. Specifically we will show how a user can authenticate at their home site and based on the attributes that are returned via Shibboleth, e.g. their different roles in the clinical trials they might be associated with, be provided with different query possibilities. These include:

- If they are a study oversight member they are able to view all data sets;
- If they are a doctor they can see summaries of the clinical data sets;
- If they are a nurse then they can only see a non-identifying subset of the data sets.

We will also show how patient consent can be captured and used to enforce access to and usage of a variety of resources. Figure 1 shows the user interface to the VOTES framework (with the nurse attribute returned and used to provide a restricted view of the data).

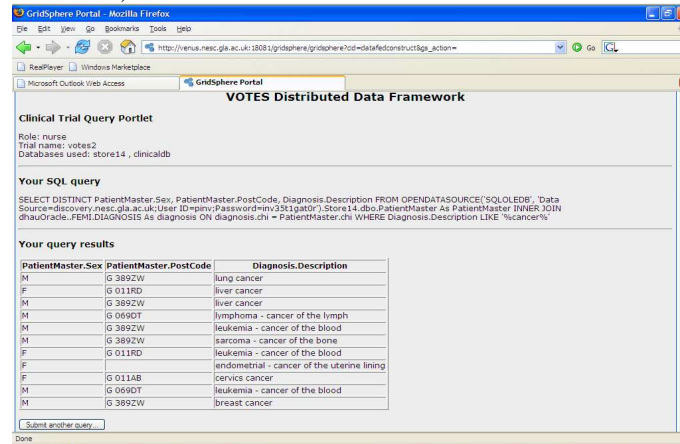


Figure 1: VOTES Distributed Framework accessed through Shibboleth

2.2 Dynamic Virtual Organisations for e-Science Education (DyVOSE) demonstration

Shibboleth can be used to provide single sign-on and authorization through recognition of remote “trusted” institutions in a federation. Shibboleth is by its nature much more static than the vision of the Grid in creating dynamic Virtual Organisations (VOs) “*on-the-fly*” where new roles, people, resources or services are added (or removed) across VOs. A key aspect in supporting such scenarios is the ability for an institution to dynamically delegate privileges to remote entities to create VO specific security attributes that will be recognized locally.

To support such scenarios, the DyVOSE project has developed a Delegation Issuing Service (DIS) (Figure 2) where a remote (and trusted) source of authority (SoA) – typically a remote sys-admin, is given the privilege to issue attribute certificates on a local security infrastructure which the remote site will recognize. In this demonstration we will show how sites wishing to establish VOs dynamically are able to create attribute certificates associated with the particular demands of the give VO via the DIS *on the fly*. 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.



Figure 2: DyVOSE Delegation Issuing Service

In demonstrating the DIS we will show a scenario implemented within the DyVOSE project where students were required to access a genome database via a PERMIS secured Grid service which would return either nucleotide or protein sequences depending upon their group (role). We will show how student roles are dynamically added to site policies by a remote SoA and used to determine which data sets are returned. These data sets are then used as part of a nucleotide or protein BLAST Grid service which utilizes the NeSC Condor pool which is also PERMIS protected. This application was undertaken by students in the advanced MSc at Glasgow as part of the Grid Computing course.

2.3 Glasgow University Early Adoption of Shibboleth (GLASS) demonstration

In this demonstration we will provide an overview of the University of Glasgow unified account management system based upon nSure technology and how it can be used to provide guaranteed authentication and authorization of university members. A key aspect of Shibboleth is trusting remote sites to authenticate and authorize individuals at their institutions, hence ensuring that an individual is from a remote site is essential (previously it was the case that Glasgow University members had numerous usernames/passwords which were not related hence an individual could leave the university and still have active accounts). In particular in this demonstration we will show how single sign on to a range of services has been achieved including: the Glasgow University e-Learning environment (Moodle) and the Glasgow University web site (webSurf) which allows students and staff to view and update central student records;

In addition, we will demonstrate how Shibboleth has been applied to provide fine grained security access to a range of brain trauma data sets for clinicians working at the Glasgow Southern General Hospital (see www.brainit.org). These data sets include MRI brain images, ECGs, as well as information on treatment and diagnosis.

2.4 Biomedical Research Informatics Delivered by Grid Enabled Services (BRIDGES) demonstration

The DTI funded BRIDGES project recently completed in December 2005. This project built data Grids using both the OGSA-DAI and IBM Information Integrator middleware. These allowed scientists to issue queries via client side tools accessible through a project portal to a database which subsequently federated these queries to a variety of public functional genomic databases such as Online Mammalian Inheritance in Man database, ensembl (rat, mouse and human genome databases), Rat Genome Database, SwissProt database, Mouse Genome Informatics database and the Gene Ontology database. In addition, BRIDGES built a compute Grid which allowed the scientists to access and run BLAST applications across a variety of resources such as all nodes of the NGS, different HPC clusters at Glasgow and Condor pools depending upon their role/privilege within the project.

This demonstration will show how Shibboleth authentication and attributes can provide simplified access to a range of computational resources including the NGS. We will show how large scale BLAST applications across all of these resources are supported as shown below where 30,000 input sequences are being compared against the aggregated *nr* (nucleotide) data sets which are pre-installed across the computational resources. (A variety of other data resources including *nr* (protein) databases and a range of microbial genome databases are also supported).

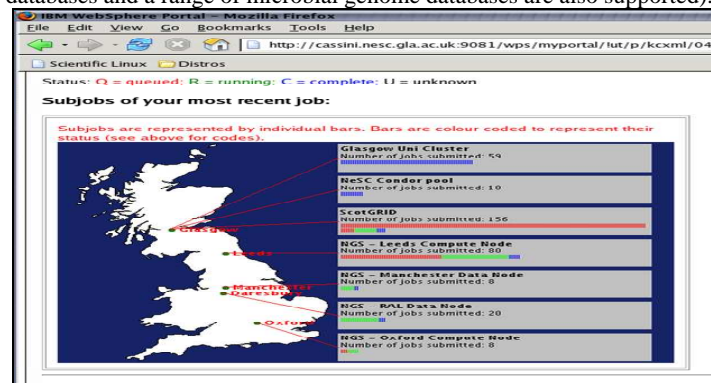


Figure 3: Shibboleth support of large scale BLASTing across NGS and other large scale computational resources with job monitoring

2.5 Meeting the Design Challenges of nanoCMOS Electronics demonstration

In this demonstration we will show how Shibboleth can be used to address licensing issues on the Grid. Although this project will start in October 2006, work has already started in using Shibboleth to provide secure access to services where licenses are needed or to protect commercial IP associated with designs, data sets and processes. We will demonstrate how security attributes for software licenses are used to restrict/enforce access to protected services.

3. Summary of Benefits of Demonstrations

This portfolio of demonstrations will show the benefits of using Shibboleth in the Grid context. A multitude of different research topics are covered from teaching to clinical trials, to bioinformatics to engineering, covering a wide range of UK funders. As such the talks and demonstrations can be tuned to the interests of the SC2006 attendees: from security focused data access, to HPC, to licensing issues, to VO mgt. In short, we won't just run one demonstration but have a portfolio of presentations and demonstrations to run across a range of research areas which we believe will be of interest to a large range of attendees at SC2006.

It is also the case that integration of Shibboleth and Grid technologies is a "hot topic" with much effort within the GGF and other places to address these issues. The DIS service in particular represents state of the art in this area with papers recently published at CCGrid 2006 and papers submitted and currently being reviewed for Grid2006. The demonstrations themselves have a high visual impact, especially those returning clinical images for the GLASS project.