Sinnott, R.O. and Ajayi, O. and Jiang, J. and Stell, A.J. and Watt, J. (2007) *User-oriented security supporting inter-disciplinary life science research across the grid.* New Generation Computing, 25 (4). pp. 339-354. ISSN 0288-3635

http://eprints.gla.ac.uk/7212/

Deposited on: 8 September 2009

User-oriented Security Supporting Inter-disciplinary Life Science Research across the Grid

Professor Richard SINNOTT, Oluwafemi AJAYI, Jipu JIANG,
Anthony STELL and Dr John WATT

National e-Science Centre, University of Glasgow
Glasgow, G12 8QQ, UK

`r.sinnott@nesc.gla.ac.uk`

Abstract   Understanding potential genetic factors in disease or development of personalised e-Health solutions require scientists to access a multitude of data and compute resources across the Internet from functional genomics resources through to epidemiological studies. The Grid paradigm provides a compelling model whereby seamless access to these resources can be achieved. However, the acceptance of Grid technologies in this domain by researchers and resource owners must satisfy particular constraints from this community - two of the most critical of these constraints being advanced security and usability. In this paper we show how the Internet2 Shibboleth technology combined with advanced authorisation infrastructures can help address these constraints. We demonstrate the viability of this approach through a selection of case studies across the complete life science spectrum.

§1   Introduction

The life science domain is booming - the explosion of research areas, exponential growth of the associated data sets and the proliferation of new discoveries across and between disciplines is unparalleled. Building infrastructures to support such a highly volatile area is a fraught process [1]. As new insights and

discoveries are made, existing research models and associated data sets have to be augmented, refined and extended to incorporate such new knowledge. This poses a fundamental challenge to infrastructure providers: how to build an infrastructure that has any form of longevity [2]?

One thing which is clear is that scientists need to be able to collaborate with one another. However scientists in the post-genomic era are wary and protective of their own research. Collaborators in grants are also potential competitors in future funding proposals. Being the first scientist to make a major discovery is a strong driver for many leading researchers both in the present and the past. Similarly the direct financial benefits from given lines of research can pay huge dividends with interest from the multi-billion dollar pharmaceutical industry given the costs they incur in developing new drugs and evaluating their effectiveness.

In this context scientists need to be ensured that they access and share trusted data from collaborators and that this is in accordance with the commonly agreed terms and goals of the collaboration. In the context of the Grid, such co-operations are often termed virtual organisations (VO). From past experiences [3] it is clear that vast amounts of the non-Grid community are uncomfortable with the Grid. System administrators regard it as a possible security threat whilst potential Grid end users are mystified and put off with the common security models needed to access and use a Grid. If the Grid is to ever gain the widespread acceptance envisaged, it needs to be as simple to use as the Internet, with client side front end tools no more complex than existing browsers.

The Shibboleth technology from the Internet2 project [4] has put forward software architecture [5] and associated protocols [6] for new models of security. Rather than a user being required to remember numerous usernames and passwords required to access resources across the internet, with the Shibboleth trust model, a user is able to access resources across a federation through signing in (authenticating) at their own local site. Through potential release and subsequent acceptance (or rejection) of security attributes by providers of services or data sets fine grained authorised access to biological data can be achieved. In this paper we present the Shibboleth model of security and outline how it has been applied across a range of biomedical projects to simplify the access and usage of Grid services and data. In all of this, a finer grained model of security is achievable.

§2   Overview of Grid Security and Shibboleth

The majority of Grid solutions today are based upon X.509 digital certificates [7] to support public key infrastructures (PKI) [8]. PKIs are based upon cryptographic technology where public and private keys are used to secure messages and information through encryption and digital signatures. The establishment of the identity of a given user (or machine) within a public key certificate is ultimately based upon trust, or more precisely trust by a community of one or more Certificate Authorities - the roots of trust chains. CAs have numerous responsibilities including issuing of certificates, issuing Certificate Revocation Lists (CRL) and they need to have well documented processes and practices which must be followed to ensure identity management.

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 RAL, a user in possession of an X.509 certificate issued by RAL can send jobs to all sites, or rather to all sites where a user has requested and been granted access to those sites. Typically with Globus based solutions [9], gatekeepers are used to ensure that signed Grid requests are valid, i.e. from known collaborators. When this is so, i.e. the DN of the requestor is in a locally stored and managed grid-mapfile, then the user is typically given access to the locally set up account as defined in the grid-mapfile.

For the existing Grid community, PKIs are a widely accepted and common model for how to support a basic level of security (authentication). The problem is however that the existing Grid community are only a small fraction of the wider e-Science and e-Research community more generally, e.g., in the UK 3500 UK e-Science X.509 certificates have been issued by RAL, but there are over 3.5 million academics across higher and further education colleges in the UK registered to use the Athens authentication infrastructure. Thus it could be claimed with some justification that the Grid has only touched the tip of the iceberg in terms of the wider research community.

There are many possible reasons for this as identified in [3]. These include the learning curve in accessing and using Grids. Most scientists do not want to gain access to a user account on a HPC resource but want instead to access a service which performs some function, e.g. BLAST in the case of the bioinformatics community. Why should a biologist go on a training course on Grid technology when all they require is access to a BLAST service on a free national HPC resource for example? Furthermore, the initial hurdles that have

to be overcome in getting on the Grid in terms of acquiring and using an X.509 certificate are non-trivial for less IT-oriented researchers. For example, users are expected to convert the certificate from their CA which they initially install in their browser into appropriate formats understandable by Grid middleware. This requires them to run obscure openSSL commands, and since openSSL is not commonly available on platforms such as Windows they are then often required to install and configure additional software. In some circumstances this is also not possible, e.g. if they do not have sufficient privileges on their PC (root access etc). In this case the researchers will instead have to refer to a local system administrator to help with the installation and configuration.

Assuming researchers have managed to obtain a certificate which they have converted into the appropriate format, they are then expected to remember necessarily strong passwords for their private keys with the recommendation to use upper and lower case alphanumeric characters. The temptation to write down such passwords is obvious and an immediate potential security weakness.

This whole process does not lend itself to the wider research community which the e-Science and Grid community needs to reach out to and engage with. Usability and addressing researcher requirements is crucial to the uptake and success of Grid technology. End user scientists require software which simplifies their daily research and not make this more complex. Given the fact that the initial user experience of the Grid currently begins with application for an e-Science certificate, this needs to be made as simple as possible, or potentially removed completely.

PKIs support authentication, however it is clear that the vast majority of researchers require much finer grained security infrastructures which support authorisation. Not just establishing the identity of a given user at a resource, but in defining and enforcing how they might access and use a given resource, or if sufficient information is not given, rejecting the request and subsequently logging the information. Despite these limitations, single sign-on is a compelling model and any refinements, extensions or new solutions for Grid security must provide similar capabilities. Such models should also be targeted to, and at the complete discretion of the resource provider to provide site autonomy. The Internet2 Shibboleth technology provides one way in which many of these issues can be resolved. Shibboleth introduces several concepts which are fundamental to access and use Grid resources. These include an Identity Provider (IdP), a Service Provider (SP) and optionally a Where Are You From Service (WAYF).
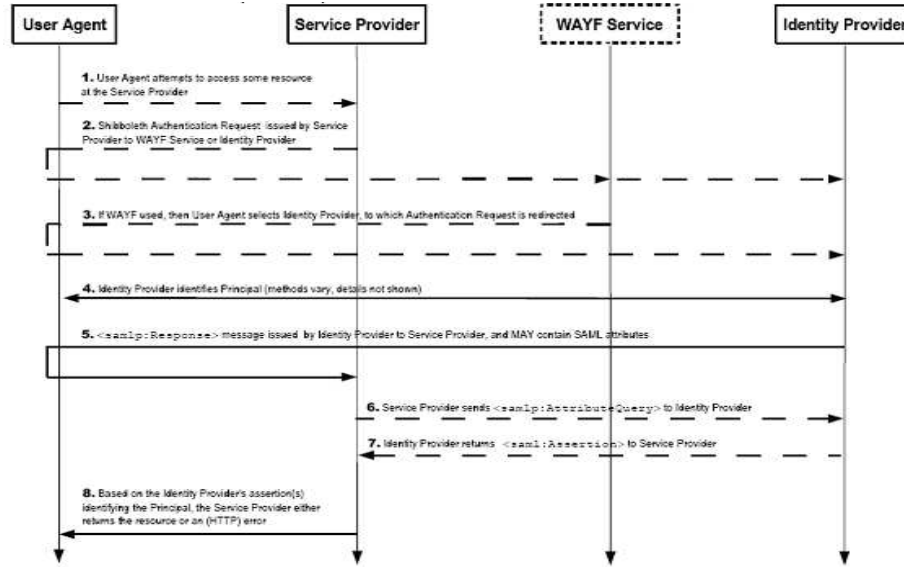
Fig. 1   Basic Shibboleth interactions for accessing a resource

The basic scenario by which these components interwork is depicted in Figure 1.

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. In the UK a single federation has been established [10]. Other international federations have also been put forward and established [11][12][13][14].

After the user has picked their home site, their browser is redirected to their sites 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 [15] authentication assertion message from the home site, asserting that the user has been successfully authenticated (or not) by a particular means using an authentication mechanism specific to the IdP. Assuming 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 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 IdPs attribute authority in a SAML attribute query message and is returned a signed SAML attribute assertion message.

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 federation provided the necessary attributes and values are presented by the IdP.

Ensuring that an institution in a Shibboleth federation can guarantee the authenticity of a user when accessing a remote resource is crucial to the overall principles upon which Shibboleth and Shibboleth federations are based. In short, institutions in a federation should trust one another. It is the case however, that users at larger institutions will likely have numerous usernames and associated passwords that are used to access a variety of services.

The directory is the part of any service which retains the authentication data, most commonly a username and a password. Until recently this information at Glasgow was closely linked to specific operating systems or infrastructures. This resulted in myriad solutions holding a variety of authentication information across the university. One of the consequences of this was that the evolution of services became tied to the platform hosting the user identities, rather than the best platform for the job. In most cases these accounts were not necessarily the same - indeed in lots of cases they were very different, and often based on a combination of central and local accounts. Thus users were expected to keep multiple accounts and multiple passwords. Under these circumstances users tended to either leave the password at the value it was when they received

it; change it to the same value as their other passwords; they have to remember multiple passwords, or they end up with passwords they cant change because changing it in one place means changing it everywhere. With multiple accounts, across multiple systems with potentially multiple different administrators coordinating changes was almost impossible. Addressing such issues is crucial for the wide scale successful deployment and take-up of Shibboleth and in trusting sites.

The above problems are not isolated. Until recently no mechanisms existed to keep the various user accounts synchronised across all of the systems used. This arrangement meant there was a high number of redundant accounts, which has meant that it was very difficult to ensure all access and privileges were removed in a timely fashion. In some circumstances users could retain rights to data and services long after they should. This was possible since different representations for the same users could in principle lead to situations where one account could be disabled, but users could retain access to services and data via a second account. A key challenge is therefore to address the whole user base since there may be no definitive source for authentication data, but rather a collection of sources.

To overcome these issues the University of Glasgow has moved to a system that offers a more consistent representation of staff and students across multiple systems that will allow management of accounts, an audit trail and the implementation of a rigorous password policy. To support this, the university has established a one to one representation between each user and their corresponding entry in the Human Resource/Registry database the definitive sources for data.

§3   Case Studies Applying Shibboleth

The vision of the Grid in seamlessly accessing and using a range of resources is a compelling one, but one that depends on supporting technologies. Single-sign on to resources is one of the fundamental requirements to the realisation of this vision. From the life science community perspective, single sign-on to a whole range of post-genomic resources (both computational and data resources) through to clinical and epidemiological data sets and services is needed. Several projects at the National e-Science Centre at the University of Glasgow have been used to explore the suitability of Shibboleth for Grid security.

3.1   Shibboleth version of the BRIDGES Grid BLAST Service

The BRIDGES project involved the National e-Science Centre at the University of Glasgow and Edinburgh with industrial involvement from IBM. The project remit was to provide a Grid infrastructure to support the Wellcome Trust funded Cardiovascular Functional Genomics (CFG) project [20] who are investigating possible genetic causes of hypertension. This consortium which involved five UK sites and one Dutch site pursued a strategy combining studies on rodent models of disease (mouse and rat) contemporaneously with studies of patients and population DNA collections.

Before BRIDGES, many of the activities that the CFG scientists undertook in performing their research were done in a time consuming and largely non-automated manner. To address this, the BRIDGES project developed a security focused data Grid using commercial [21] and open source Grid middleware [22]. Information on the data Grid that was developed within BRIDGES is described in [23][24][25]. In undertaking their research the CFG scientists also required simple access to large scale HPC resources, to run compute intensive bioinformatics applications such as Basic Local Alignment Search Tool (BLAST) [26].

There are several existing BLAST implementations that take a variety of approaches [31] - [36] but none of these suited the particular requirements of the BRIDGES project. After researching the different methods available, the final design adopted in BRIDGES implemented parallelization of BLAST at the level of the input data.

To simplify the user experience in accessing and using large scale HPC resources, it was decided to remove digital certificates from the end user environment and replace them with simple username and password authentication at a central project web portal. The model assumed throughout the lifetime of BRIDGES was that the end users would only have a web browser through which they would access and use the Grid resources.

BRIDGES supported a fine grained security infrastructure based upon the Privilege and Role Management Infrastructure Validation Standard (PERMIS) (www.permis.org), whereby distinctions between different privileged and non-privileged users were defined and subsequently enforced - implemented as a hierarchy of access (in decreasing order of access) to NGS resources, Glasgow HPC resources and the local NeSC Condor pool.
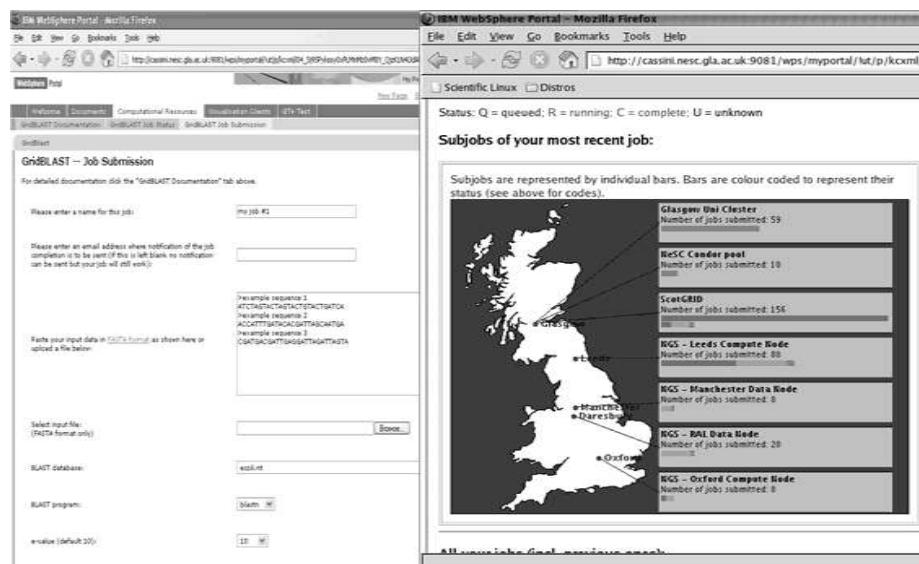
Fig. 2   Grid BLAST front end

The Shibboleth version of the Grid BLAST service did not require users to log in to the project portal. Instead when users directed their browsers to the project portal for the first time, i.e. when no security context had been established, they were automatically redirected to the WAYF service. They would then select their own institution from those that were listed.

Once redirected to their home IdP, which at Glasgow was a local LDAP server, users would log-in with their own usernames and passwords. Once authenticated, the attributes that were returned to the SP were used to enforce subsequent authorisation decisions by the service. We note that whilst BRIDGES VO specific attributes could be defined and returned, the DN of the user from the IdP was sufficient to allow an authorisation decision to be made. Once successfully signed in the front end to the Grid BLAST service is accessible as shown on the left of in Figure 2. This provides access to a range of genomic and microbial data resources which can be BLASTed against. The service supports protein and nucleotide sequence searches and allows upload of input sequences or cut/paste of sequences in FASTA format. To support large scale BLASTing users can select options to be emailed the results when the jobs are completed, or they can interactively see the status of the jobs across the various Grid resources (whether they are queued, completed, running) shown in the right of Figure 2.

We note that this model of applying Shibboleth where the user identity (given by the Distinguished Name) is returned and subsequently used to make authorisation decisions, raises issues in the application of Shibboleth. Shibboleth has been developed to support user anonymisation and privacy in accessing and using resources across a federation. However, with the Grid model, knowing which user is accessing a resource, especially in the biomedical domain is crucial. We also note that whilst Shibboleth supports user anonymisation and privacy it is not mandatory and information such as the DN of the user from an IdP to an SP can be returned. The policies on what information and attributes an SP can ask for and what information an IdP is prepared to release will form part of the overall federation contract. There is no obligation on an IdP releasing potentially sensitive information about a given user. However if an SP requests certain attributes to be returned for example which the IdP refuses to release then the SP is completely free to refuse to grant access to their own resource. SP autonomy is thus assured.

As stated, the primary benefit of the Shibboleth enabled version of the BRIDGES Grid BLAST service is that the user no longer needs to remember a username and password for a given portal, and only for their home site. Whilst only needing to know a single username and password is a key benefit in applying Shibboleth, the true benefits arise when the user wishes to access a multitude of different resources across many sites, as is typically the case in supporting systems-biology based research from the genotype to the phenotype and population studies, i.e. supporting single sign-on across a range of resources and sites. Whilst BRIDGES was targeted to the genetic end of the spectrum, the VOTES project was targeted towards the other extreme and was focused on clinical trials and epidemiological studies.

3.2 Shibboleth version of the VOTES Data Federation Framework

The VOTES project was funded by the MRC for 3-years and began in October 2005. The project involves the universities of Glasgow, Oxford, Imperial College London, Nottingham, Leicester and Manchester. The overall goal of VOTES is to develop a Grid framework through which a multitude of clinical trials and epidemiological studies can be supported. Thus rather than engineering bespoke solutions for a given trial or study, VOTES intends to provide an infrastructure where a multitude of trials and studies can be developed and sup-

ported, each with their own particular nuances in terms of the data that is being accessed, the security policies that apply etc.

Two of the key processes at the heart of clinical trials has been identified as patient recruitment and data collection. Throughout these processes it is essential that the study or trial is conducted according to a strictly defined protocol. This will focus on what information is being collected, for what purpose and how it will subsequently be used both within and following the trial. A key element of this is ensuring that the different people with different roles within the trial can only access and use the different data sets and services associated with their particular role in the trial.

The Grid infrastructure developed within VOTES has been described in [27], [28] and [29] - the current implementation combining access to and usage of a range of software and data sets in widespread use across the NHS in Scotland.

The project has defined and implemented a fine grained authorisation infrastructure based upon an access matrix. In this model, very fine grained security is achievable that allows a clinical trials co-ordinator to define access to and usage of individual tables, rows and columns across the range of distributed clinical data sets and attach these with trials specific roles.

The Shibboleth version of this infrastructure follows a similar access and usage pattern as described above in section 3.1 for BRIDGES. Namely, that the user attempts to access the VOTES portal and is initially redirected to the WAYF service where they select their home IdP. We note that when a user has signed in already, e.g. they have authenticated themselves to access the BRIDGES portal, provided the user is using the same browser, they will in principle, automatically be allowed access to the VOTES portal. Here the term in principle depends on whether the attributes necessary for access to the VOTES portal were already released to the BRIDGES portal. In Figure 4 the portlet is shown which details the attributes that are returned from the IdP. As can be seen one of the roles is investigator (the other roles are used for a different demonstration in the education domain). The common name (CN) is also returned (CN=Guest2) along with information on the Shibboleth origin (another name for the IdP). Based upon this authentication information and the attributes returned, the portal will ensure that the authenticated user is restricted to seeing the trials and associated data associated with their particular role (privilege).

We note that through the attributes presented by Shibboleth, the user
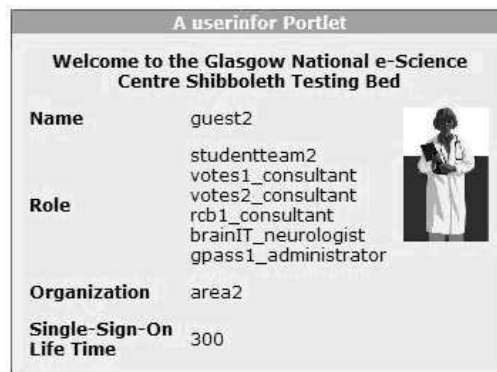
Fig. 3 Authenticated user attributes delivered via Shibboleth to the VOTES portal

is effectively restricted to a given view of the distributed clinical data sets. The views themselves are statically assigned during the establishment of a given clinical trial. Thus it is not possible (nor ever likely) that arbitrary queries can be run across clinical data sets containing identifying patient data. To support the process of aligning user views of data with the roles and authorisation policies, tools (JSR-168 compliant portlets) are being developed for use by non-Grid savvy clinicians will ultimately drive the whole process of deciding what data is made available to a given role in a given trial. This is possible through a detailed knowledge of the data schemas used for holding the clinical data sets in Scotland. Details of these are given in [27][28][29].

3.3   Shibboleth enabled Neurological Research

The Institute of Neurological Sciences at Southern General Hospital in Glasgow is the co-ordinating centre for brain trauma research across Europe (www.brainit.org). A central repository for brain trauma data has been established. Access to and usage of the brain trauma data sets contained in this repository is strictly controlled with a variety of fine grained security policies in place. Information and data on over 400 brain trauma patients has been added to the BrainIT central repository. These data comprise various neurological DICOM imaging data sets of brain trauma patients across Europe along with physiological data sets related to their monitoring and current course of treatments.

Collecting and sharing these data sets allows neurological researchers to better understand the way in which brain trauma patients are treated and
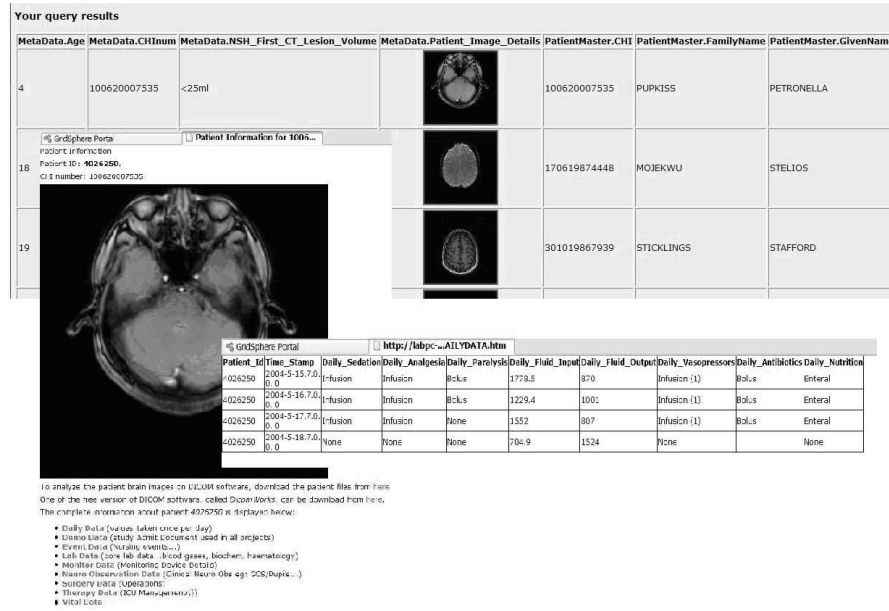
Fig. 4 Neurological data returned from the BrainIT portal for privileged user

effectiveness of treatment courses in the large. However, pertinent questions and comparisons related to the real time care and treatment of individual patients can also be achieved.

Access and use of the BrainIT resources demands appropriate security policies are upheld. These security policies include a variety of conditions that mainly focus on the need for a user to have contributed to the data sets, or for a user to hold a tenure of authority in the appropriate domain before they can view the data sets.

A Shibboleth-enabled version of the BrainIT repository has been developed as part of the GLASS project. In this scenario, various roles specific to the BrainIT data resource have been defined and are used by the BrainIT portal to enforce local authorisation decisions. As with access to and usage of the BRIDGES and VOTES resources via Shibboleth, single sign-on to the Grid-enabled BrainIt resource is directly supported without further requirements to authenticate. Figure 4 shows the results of a privileged user accessing and using the BrainIT - the patient data is at the back, the close-up of the DICOM image is returned when the original image is clicked upon, and a sample of the in-depth monitoring and physiological data for this patient can be seen on top.

§4   Conclusions

The life scientists interest in tackling the big questions like how does a brain work, why do people who eat less tend to live longer, what genes cause cancer, etc cannot work in isolation. They need to access and use a wide range of information much of which may well have very fine grained security associated with it. The Grid provides an infrastructure whereby heterogeneous, highly distributed data sets can be accessed and integrated. A simple user oriented model of security is essential that should be cognisant of existing security infrastructures and policies, e.g. how local authentication is supported. Shibboleth provides for many direct advantages in this domain: it supports single sign-on to a widespread range of resources across a federation; it recognises and respects local security infrastructures policies and procedures; it supports users only having to know their own local institutional usernames and passwords; it allows for very fine grained authorisation infrastructures to be supported based on returned attribute sets.

Shibboleth on its own offers a largely static mechanism whereby the roles and attributes needed to access resources have to be defined and agreed in advance. This does not support the more dynamic Grid model where new virtual organisations need to be created dynamically and require richer attribute sets, specific to the different roles across a given project. Our next plans are to consider how this service can be applied in this domain for dynamic attribute creation and recognition across clinical domain boundaries. This will overcome the largely static model of security upon which Shibboleth federations are based. Extending this to support more dynamic VO-specific roles and attributes is essential.

Ultimately Grid or Shibboleth based technological solutions have to be supported by a willingness to collaborate by the biomedical research community. To support this, a better model for BrainIT and for life science researchers more generally is to maintain their own data sets and provide secure (federated) access to them. A single centralised repository is a single source of failure which may ultimately have fatal consequences. Resilience through distribution and replication is one obvious advantage through a Grid based approach, however a further key benefit is the fact that researchers are in control of their data. As noted previously the biomedical research community are wary of others accessing and using their data before they have exploited them fully and published results from them in respected journals for example. Defining and enforcing local access con-

trol to collaborating researchers with local policies is an important psychological process which will encourage data sharing more generally.

Finally the key benefits of Shibboleth are to provide single sign-on to a range of resources. The scenarios presented in this paper and being extended to include non-Grid based resources such as the University of Glasgow Virtual Learning Environment (Moodle) amongst many other resources. Similarly, Shibboleth versions of services for secure access to and comparison of microarray expression profiles is also being undertaken within the Grid Enabled Microarray Expression Profile Search (GEMEPS) project at NeSC in Glasgow [30]. This will allow researchers to find and compare similar experiments from a variety of perspectives: based upon the same platform, e.g. GPL570; the same species; the same disease; or through more advanced queries based upon the genes that are expressed in experiments and their levels of expression.

References

1) R.O. Sinnott, M. Bayer, "Controlling the chaos: Developing Post-Genomic Grid Infrastructures," in *Proceedings of the Life Science Grid Conference (LSGrid 2005)*

2) R.O. Sinnott, P. Lord, A. McDonald, "Large scale data sharing in the life sciences: Data standards, incentives, barriers and funding models (The Joint Data Standards Study)," *prepared for the Biotechnology and Biological Sciences Research Council, the Department of Trade and Industry, Joint Information Systems Committee, The Medical Research Council, The Natural Environment Research Council and The Wellcome Trust*

3) R.O. Sinnott, "Grid Security: Middleware, Practices and Outlook" *Prepared for the Joint Information Systems Committee for Support for Research, November 2005*

4) "Internet2 Shibboleth Technology," *http://shibboleth.internet2.edu*

5) "Shibboleth Architecture Technical Overview," *http://shibboleth.internet2.edu/docs/draft-mace-shibboleth-tech-overview-latest.pdf*

6) "Shibboleth Architecture Protocols and Profiles," *http://shibboleth.internet2.edu/docs/draft-mace-shibboleth-arch-protocols-latest.pdf*

7) "ITU-T Recommendation X.509 (2001) — ISO/IEC 9594-8: 2001, Information Technology," *Open Systems Interconnection - Public-Key and Attribute Certificate Frameworks*

8) R. Housley, T. Polk "Planning for PKI: Best Practices Guide for Deploying Public Key Infrastructures" *Wiley Computer Publishing, 2001*

9) "The Globus Toolkit" *http://www.globus.org*

10) "UK Shibboleth Federation" *http://www.sdss.ac.uk*

11) "Swiss SwitchAAI Federation" *http://www.switch.ch/aai*

12) "Finnish HAKA Federation" *http://www.csc.fi/suomi/funet/middleware/english*

13) "Australian Meta Access Management System" *https://mams.melcoe.mq.edu.au/zope/mams/kb/shibboleth*

14) "US InCommon Federation" *http://www.incommonfederation.org"*

15) "Security Assertion Markup Language (SAML) version 2.0, March 2005" *http://www.oasis-open.org/specs/index.php*

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/*

17) "DTI funded Biomedical Research Informatics Delivered by Grid Enabled Service (BRIDGES) project" *www.nesc.ac.uk/hub/projects/bridges*

18) "MRC funded Virtual Organisations for Trials and Epidemiological Studies (VOTES) project" *www.nesc.ac.uk/hub/projects/votes*

19) "JISC funded Glasgow university early adoption of Shibboleth (GLASS) project" *www.nesc.ac.uk/hub/projects/glass*

20) "Cardiovascular Functional Genomics (CFG) project" *www.brc.dcs.gla.ac.uk/projects/cfg*

21) "IBM Information Integrator" *http://www-306.ibm.com/software/data/DL*

22) "Open Grid Service Architecture Data Access and Integration Two (OGSA-DAIT)" *www.ogsadai.org*

23) R. O. Sinnott, M. M. Bayer, J. Koetsier, A. J. Stell "Grid Infrastructures for Secure Access to and Use of Bioinformatics Data: Experiences from the BRIDGES Project" in *1st International Conference on Availability, Reliability and Security, (ARES06), Vienna, Austria, April, 2006*

24) R.O. Sinnott, D. Houghton "Comparison of Data Access and Integration Technologies in the Life Science Domain" in *Proceedings of UK e-Science All Hands Meeting, September 2005, Nottingham, England*

25) R. O. Sinnott, M. M. Bayer, J. Koetsier, A. J. Stell "Advanced Security on Grid-Enabled Biomedical Services" *Proceedings of UK e-Science All Hands Meeting, September 2005, Nottingham, England*

26) R.O. Sinnott, M. Bayer "Distributed BLAST in a Grid Computing Context" in *Proceedings of First International Workshop on Distributed Data Mining in Life Science, Konstanz, Germany, September 2005*

27) A.J. Stell, R.O. Sinnott, O. Ajayi "Secure Federated Data Retrieval in Clinical Trials" in *Telemedicine 2006 conference, Banff, Canada, July 2006*

28) R.O. Sinnott, A.J. Stell, O. Ajayi "Development of Grid Frameworks for Clinical Trials and Epidemiological Studies" in *HealthGrid 2006 conference, Valencia, Spain, June 2006*

29) A.J. Stell, R.O. Sinnott, O. Ajayi "Supporting the Clinical Trial Recruitment Process through the Grid" in *Nottingham UK e-Science All Hands Meeting, September 2006*

30) "BBSRC funded Grid Enabled Microarray Expression Profile Search (GEMEPS) project" *www.nesc.ac.uk/hub/projects/gemeps*

31) Pedretti K.T., Casavant T.L., Braun R.C., Scheetz T.E., Birkett C.L., Roberts C.A. "Three Complementary Approaches to Parallelization of Local BLAST Service on Workstation Clusters in *Proceedings of the 5th International Conference on Parallel Computing Technologies, p.271-282, September 06-10, 1999*

32)   Braun R.C., Pedretti K.T., Casavant T.L., Scheetz T.E., Birkett C.L., Roberts C.A.   "Parallelization of local BLAST service on workstation clusters."   in *Future Generation Computer Systems 17: 745-754, 2001*

33)   Clifford. R., Mackey A.J.   "Disperse: a simple and efficient approach to parallel database searching"   in *Bioinformatics 16: 564-565, 2000*

34)   Mathog D.R.   "Parallel BLAST on split databases"   in *Bioinformatics 19: 1865-1866, 2003*

35)   Darling A.E., Carey L., Feng W.   "The design, implementation and evaluation of mpiBLAST."   in (*Proceedings of ClusterWorld Conference and Expo and the 4th International Conference on Linux Clusters: The HPC Revolution, 2003*

36)   Hokamp K., Shields D.C., Wolfe K.H., Caffrey D.R.   "Wrapping up BLAST and other applications for use on Unix clusters"   in Bioinformatics 19: 441-442, 2003