



JISC Final Report

Project Information			
Project Identifier	<i>To be completed by JISC</i>		
Project Title	A Proxy Credential Auditing Infrastructure for the UK e-Science National Grid Service		
Project Hashtag			
Start Date	1 st Jan 2010	End Date	31 st March 2011
Lead Institution	Thames Valley University		
Project Director	Prof. Richard Sinnott		
Project Manager	Dr. Wei Jie		
Contact email	wei.jie@tvu.ac.uk richard.sinnott@glasgow.ac.uk		
Partner Institutions	National e-Science Centre, University of Glasgow		
Project Web URL	http://www.nesc.ac.uk/hub/projects/pca		
Programme Name	Access and Identity Management Programme		
Programme Manager	Christopher Brown		

Document Information			
Author(s)	Wei Jie and Richard Sinnott		
Project Role(s)	Project Manager, Project Director		
Date	May 2011	Filename	
URL			
Access	This report is for general dissemination		

Document History		
Version	Date	Comments
1.0	12 th May 2011	Draft created by Wei Jie
1.1	16 th May 2011	Revised by Richard Sinnott
1.2	20 th May 2011	Revise by Christopher Bayliss

Table of Contents

1	ACKNOWLEDGEMENTS	3
2	PROJECT SUMMARY	3
3	MAIN BODY OF REPORT	3
3.1	PROJECT OUTPUTS AND OUTCOMES	3
3.2	HOW DID YOU GO ABOUT ACHIEVING YOUR OUTPUTS / OUTCOMES?	4
3.3	WHAT DID YOU LEARN?	5
3.4	IMMEDIATE IMPACT.....	6
3.5	FUTURE IMPACT	7
4	CONCLUSIONS	7
5	RECOMMENDATIONS	8
6	IMPLICATIONS FOR THE FUTURE	8
7	REFERENCES	9
8	APPENDICES (OPTIONAL)	10

1 Acknowledgements

This project, a Proxy Credential Auditing Infrastructure for the UK e-Science National Grid Service (PCA), was funded by JISC under the Access and Identity Management Programme. It was led by Thames Valley University in collaboration with the National e-Science Centre at the University of Glasgow. We would like to give special thanks to JISC programme manager, Christopher Brown, for his timely guidance and support. We would also like to thank our project partners for their involvement and advice during the project.

2 Project Summary

Single sign-on and delegation of privileges are fundamental tenets upon which e-Infrastructures and Grid-based research more generally have been based. The realisation of single sign-on and delegation of privileges in accessing resources such as the UK e-Science National Grid Service (NGS) is typically facilitated by X.509-based Public Key Infrastructures (PKI) and exploitation of proxy certificates. This model can be categorised by authentication-oriented access and usage of resources. It is the case however that proxy certificates can potentially be obtained and abused by a malicious third party without the knowledge of the holder. In this project we proposed a novel proxy credential auditing solution that addresses this issue directly. We have designed and developed a prototype proxy credential auditing infrastructure and demonstrated this working in widely distributed and heterogeneous research environments exemplified by the NGS. The proxy credential auditing infrastructure has been integrated into the NGS. Testing of the proxy credential auditing service was undertaken with case studies from the UK Economic and Social Research Council (ESRC) funded Data management through e-Social Science (DAMES – <http://www.dames.org.uk>) project.

We would recommend making this auditing service available to the NGS for longer term auditing and monitoring of its customer and research base. It is also expected to demonstrate the use of this auditing service in international settings as international auditing efforts represent a key requirement in establishing global Grid infrastructures. In addition more case studies of the proxy credential auditing service need to be developed to demonstrate a broader application. Finally, with the wider deployments of the proxy credential auditing infrastructure, it will also make it possible to explore further research issues, in particular, identifying irregular patterns of proxy certificate usages. In this regard, training algorithms such as Bayesian Neural Network could be adopted.

3 Main Body of Report

3.1 Project Outputs and Outcomes

Output / Outcome Type <i>(e.g. report, publication, software, knowledge built)</i>	Brief Description and URLs (where applicable)
A proxy credential auditing infrastructure middleware and overall methodology to support proxy credential auditing and monitoring	The project has presented a novel proxy auditing solution. The auditing service developed through the project can track the usage of proxy certificates, and capture and analyse proxy certificate usage trends (patterns). The software developed has been well documented.
Integration of a proxy credential auditing infrastructure into the NGS, and implementation of a range of relevant case studies	The proxy credential auditing service has been deployed at test NGS sites including ScotGrid. Example applications such as DAMES to use the auditing service have been developed and demonstrated.

Project workshops for community engagement and technology transfer	Through workshops that were held, a range of scenarios and demonstrations of the auditing service were given and feedback collected from key security experts from the NGS. These were then used for further development and refinement of the proxy credential auditing infrastructure.
Project publications (as shown on the right column)	<ul style="list-style-type: none"> • Christopher Bayliss, Richard Sinnott, Wei Jie and Junaid Arshad. <i>A Proxy Credential Auditing Infrastructure for UK e-Science National Grid Service</i>, a poster published on the 11th ACM/IEEE International Conference on Grid Computing (Grid 2010), Belgium, October 2010. • Christopher Bayliss, Richard Sinnott, Wei Jie and Junaid Arshad. <i>The Design, Development and Application of a Proxy Credential Auditing Infrastructure for Collaborative Research</i>, a research paper published on the 5th International Multidisciplinary Conference on e-Technologies (MCETECH 2011), Switzerland, January 2011. • Christopher Bayliss and Richard Sinnott. <i>Auditing Credential Usage in Grid Environments</i>, New Zealand eResearch Symposium, Dunedin, New Zealand, June 2011. • Christopher Bayliss, Richard Sinnott, Wei Jie and Junaid Arshad. <i>Auditing Usage of Collaborative Resources</i>, in progress for the UK e-Science All Hands Meeting, York, September 2011.

3.2 How did you go about achieving your outputs / outcomes?

All through the project we received sufficient support from the JISC program, and participating projects. This ensured that the problems we encountered were resolved in a timely manner, and that the project kept going as scheduled.

The project itself has been a collaboration between TVU and NeSC Glasgow involving a core project team of four members: a project director, a project manager, one RA at NeSC and one RA at TVU. The project team drew upon a research portfolio across a broad range of issues in Grid security. In particular, numerous past/on-going projects are directly relevant to this project and our extensive knowledge and experience have been fully exploited in this project. The project team members have also collaborated previously through a range of projects and publications. This experience helped to ensure that the work was conducted in the most efficient manner.

Work packages were clearly defined and distributed to appropriate team members at the very beginning of the project. The project subsequently carried out two phases of work, each of which had associated deliverables demonstrating the progress made. In the first phase (start Jan 2010 – end Jun 2010), we established a Globus Security Infrastructure (GSI)-audit enabled proxy credential auditing service on a test infrastructure. We also developed/extended a secure web service including associate back-end database and proxy credential audit record structures. These were used to conduct initial testing of the feasibility of the work as a whole and to illustrate the potential of the technology to the NGS at targeted workshops. In the second phase (start Jul 2010 – end Mar 2011), the auditing service was integrated into a test NGS site and demonstrated in projects and application domains, focusing on the monitoring and training of proxy usage. During the project milestones and deliverables were checked according to the set schedule. Feedback and refinements to the software developed have been continuous throughout the project.

On the technical development, we attempted to re-use and extend existing solutions wherever possible. Initially our focus was on a Globus Incubator project – Grid Proxy Auditing Infrastructure, however we quickly found that this technology would not address the specific needs of the NGS. Specifically this software was based on a new implementation of an audit-enabled enhancement to GSI. As a result, whilst supporting the basic auditing capabilities, the work described there had issues in its widespread deployment. Most importantly, it required development and roll-out of a new version

of GSI to resource providers such as the NGS. There are numerous pragmatic aspects which make this non-trivial to achieve and other models were thus required. Furthermore the design used could only record successful authentications. While it would have been possible to add support for rejected credentials recording any other form of action would have required an additional system.

An improved model of auditing is to provide a transparent auditing layer to the GSI software, i.e. compatible with existing GSI security models on the NGS but allows auditing and monitoring to be supported. This is the approach that has been taken in this work. Furthermore, it was an implicit requirement that the audit-enabled version of GSI we developed, would not adversely impact upon the performance of the overall system from both a resource provider and an end user perspective. Here performance includes both real time performance for message processing and ultimately, the stability and robustness of the software itself. Therefore, there was the need to integrate with existing NGS supported software stacks with as little impact as possible. These requirements directly shaped the project work.

A major issue that arose during the later phases of the project was the changing standards and their implementation. The entire system was developed in Ruby using the JRuby implementation of the language. The use of JRuby allowed us the opportunity to embed both Java or shared libraries should we have required it. An abandoned, early approach used Globus' GSI library directly allowing the security stack to remain unmodified. However this proved to be too fragile and difficult to debug stalling any attempt to integrate cleanly with the Globus in this manner.

Information gleaned from these forays into the Globus code base suggested that scripts could be inserted between Globus components. This method proved to be more viable as much of the information required is available but very little of it is documented requiring significant trial and error to properly identify everything that was present.

With a method of extracting data identified we needed a mechanism to collate data from within the system and forward auditing information as appropriate. To fulfill this requirement we built a Ruby on Rails web application. Ruby on Rails is an ideal choice here as it provides a REST interface by default making receiving audit events relatively simple.

3.3 What did you learn?

In this project the proposed architecture of the proxy auditing service was based on the extension of the mainstream Grid middleware, i.e. Globus toolkit. However, we found there were several technical issues and complexities when dealing with the Globus software. It remains poorly documented and writing code to integrate with it is a laborious process (however this was achieved!). The project in particular required a very low level understanding of the GSI component of the Globus software and the handshaking mechanisms that exist there – again this was achieved but at the cost of time and effort. It was also discovered that GSI does not encrypt messages by default and is thus itself not secure unless the both client and server explicitly enable it. It was further determined that the Globus GRAM client does not perform this by default nor does it provide the option for the user to enable wire encryption. In this work, these issues were not insurmountable, but they have required a huge investment of time and expertise in understanding the poorly documented source code. Having said this, the technical expertise that now exists in the lower level implementation details of Globus and X.509 PKI-based systems has now been garnered and will be applied in numerous on-going projects.

The project also had dependencies on the NGS themselves which due to resource (human) limitations, introduced delays. In particular the need for the test cases illustrating the Workload Management System (WMS) of the NGS caused significant delays. This component was crucial to the work as a whole as it provides a level of transparency at the heart of the Grid, i.e. it allows a user to simply submit jobs where the resources that are to be used to execute the jobs are selected at run time by a software component (the WMS) and not by the users themselves. This in principle should

have been a straightforward task, but caused significant delays to the full blown NGS-wide testing of the proxy credential auditing software.

In the course of this project, it was planned to exploit the proxy credential auditing software in other projects, e.g. the Engineering and Physical Sciences Research Council (EPSRC) funded nanoCMOS project. However, after much development and effort in delivering a large scale security-oriented e-Infrastructure for nanoCMOS, it was decided that many of the core data sets (commercial design libraries) and their use in associated simulations, could simply not run on public resources such as the NGS. Due to commercial and legal considerations, the only resources that could be used for this purpose was a HPC cluster inside of the Department of Electronics and Electrical Engineering at the University of Glasgow. As an aside, it took two years of legal negotiations with ARM to obtain a restricted license model to obtain the transistor design libraries used in commercial offerings. It was possible to run exemplar systems with dummy data and libraries on the NGS, but the core focus of nanoCMOS was on production-level systems using actual data. The principles of proxy credential usage as demonstrated through the DAMES project remain unchanged however. Anecdotally this is a key lesson on the vision of what can be done technically, and the commercial and legal realities of doing it for real on resources such as the NGS.

We repeatedly encountered problems with the Globus Toolkit. Much of it is un- or inaccurately documented requiring the use of source code inspection, debuggers, packet traces and test code to begin to pick apart how it works. In many areas it appears that what was, or was intended to be, an interface between components has atrophied to the point where separate components are now, in essence, tightly coupled. Given that most of Globus is written in C the time cost of exploratory delves into the code base are high.

GRAM and GSI are very close to being HTTP 1.1 over TLS with the main difference being the lack of a mechanism for delegating credentials. Had this been this case this project would have been significantly easier as passing HTTP requests through a series of filters is a solved problem with many mature tool kits available. Performing the same task in Globus, as mentioned previously, is significantly more difficult.

While support for an auditing mechanism on the NGS is wide spread agreement on what such a system would look like varies. The timeliness of audit messages is one of the simplest points of contention. Should audit messages arrive in near real time, on the order of a few seconds to a few minutes, in order to monitor the system in flight or should they only need to arrive eventually, if at all, to be used only in postmortem analysis. Many other features were suggested over the course of the project.

As our system permitted the actions a credential was used to perform once authenticated to be recorded we realised that this would likely produce a torrent of meaningless events as actions from across the worker nodes of a service reported actions. To remedy this we introduced the idea of actions having causal parents. Therefore sequences of events could be formed into graphs allowing the provenance of an action to be shown.

3.4 Immediate Impact

Security underpins collaborative research as supported through the Grid and e-Infrastructures. Auditing the usage of proxy credentials offers a direct way to minimise the threats for Grid security and represents a significant step forward for resource providers and end users alike. The achieved outputs of our project are twofold: the auditing and monitoring software itself and the overall methodology used to support proxy credential auditing and monitoring and identification of potential misuse. This work has also lead to significant new research areas, e.g. a PhD exploiting the outputs of this project is underway at the NeSC in Glasgow (The PhD student has recently completed the first year examination of her work building on proxy credential usage patterns as the basis for potential credential misuse).

Our work is immediately beneficial to the NGS and enables them to provide a service to HEIs and other resource providers that allows monitoring of the use of proxy certificates, thus providing a more secure and trusted environment for researchers to collaborate within. In addition this work has been of great interest to a number of Grid computing and e-science projects both nationally and internationally.

3.5 Future Impact

We continue to have discussions with the NGS on the proxy credential auditing service and how it might best be sustained as a key part of the NGS for longer term auditing and monitoring purposes of its customer and research base. We will continuously update the NGS technical board and their associated personnel on the work as a whole include the benefits for its wider adoption and practical issues for its wider deployment. We also aim to demonstrate use of this auditing service in a wider national and international setting including use of ScotGrid, TeraGrid in US, D-Grid in Germany, etc. Our project might also directly shape cross-country efforts such as the European Grid Infrastructure (EGI).

However it is the case that future of many of these efforts remains in a state of flux with different funding streams coming to an end, and other technologies now being pushed forward. Thus the considerable focus on Cloud technology and its use across the NGS is something that has now come to the fore. We believe that this work could directly shape these kinds of developments however our focus has been predominantly based around the Globus software stack and its use on the NGS. Even here however, the technologies themselves are evolving with Globus Toolkit 5 (GT5) now available with its GRAM-like interfaces and functionality. Whilst we have not directly undertaken work to ensure that the proxy credential auditing software developed here is compliant with GT5, we believe that this should be the case. We fully recognise the evolving nature of the software systems and their usage is an on-going effort though and one that is only worthy of following once it has been established that GT5 (as an example) is robust and will be deployed more widely across the NGS. History has shown that the availability of solutions GT2-GT3-GT4-GT5 does not always equate with their acceptance, deployment and usage by resource providers and end users.

Discussions are also on-going with a range of international projects and communities of the work as a whole, e.g. interest has arisen in South Korea (following keynote talks given there) and in Australia where Prof. Richard Sinnott is now based. We intend to actively continue to build upon the work as a whole and ensure it is embedded as widely as possible

4 Conclusions

In this project we have designed and developed a prototype of proxy credential auditing infrastructure and demonstrated its usage for tracking credentials across distributed research infrastructures like the NGS. This proxy credential auditing infrastructure has been demonstrated to and accepted by the NGS technical security staff and subsequently integrated into the NGS. We have shown the practical demonstration of this service with example case studies taken from the DAMES project. More details on this work and the technical aspects of the project as a whole are given in the Appendix.

As identified above, we continue to work in an evolving climate where funding for infrastructure such as the NGS and similar HPC facilities remains uncertain. The technological landscape is also evolving with new versions of Globus Toolkit now available, and much more focus on Clouds and the role of virtualisation to support research and research communities. Whilst we believe that the auditing infrastructure and the approach taken is a generic one, we also recognise that the direct transfer of software systems into these environments will no doubt throw up new software requirements for tuning of the system. However, we also believe that the technical experiences in developing and delivering the proxy credential auditing software could directly shape these future efforts. We are also

keen to ensure that the reinventing of the wheel, whether it is through Cloud based solutions or other, is avoided. It remains to be determined how achievable this actually is.

5 Recommendations

Several recommendations can be made to this work as a whole, which in part are based on the bigger picture of e-Science in the UK. We have demonstrated this working system to the NGS and continue to work with them on the transfer and deployment of this to their community. Locally in Glasgow, the ScotGrid resource has successfully supported this capability.

The technological landscape is changing however and adaptations and refinements to this work will be needed in the future, e.g. to support Cloud-based security models and tracking of inter-Cloud collaborative usage where public, private and hybrid Clouds co-exist and serve different communities. Such follow on projects are required if this work is to be transferred and avoid reinventing of wheels

We also recommend further education and take up of “best security practice” to the wider community. The software developed here is just one part of the security story and community (users and resource providers) need to be aware of the do’s and don’ts of security – both technical and non-technical. The UK e-Science Security Task Force existed for this purpose but for whatever reason, this effort gradually faded away and no core team has since been targeted with supplying this expertise or guiding the community at large. For the proxy credential auditing software to be optimally used, such a body along with guidance on its usage and ways to identify misuse should be established.

We recognise that the software alone will not tackle site security issues for example, but needs all providers to work collectively on this, e.g. NGS core nodes, partner nodes, affiliates etc. We are now in a situation where we can track proxy credential usage, however this work is still in its infancy and the algorithms we have developed (and continue to develop) depend on training of software systems to predict potential credential misuse. There has been little mainstream attempt at recording individual or virtual organisation level patterns of usage of Grid resources through proxy credentials. Without this, the training data sets required for the algorithm prediction engines will remain immature.

Technology transfer of project results is a key component of the work as a whole. We continue to publish papers in this space and have plans to submit papers to the UK e-Science All Hands Meeting and related conferences for example.

There is much work in the area of international alignment of access management federations using technologies such as Shibboleth. We see this work as highly complementary and something that should be followed. If truly international collaborative research exploiting Grid resources is to be achieved, then alignment of this work and Shibboleth federations to access Grid resources is urgently needed. In addition more case studies of the proxy credential auditing service need to be developed to demonstrate a broader application. The DAMES exemplar of running R-scripts on the NGS for data processing is one example but the issues of long term credential use and the need to deal with scenarios where credentials are refreshed when they are due to expire will open up new challenges that need to be addressed, i.e. ensuring that proxy credentials are not continually refreshed by masqueraders for example.

6 Implications for the future

This work has allowed exploration and development of detailed security solutions that would not have been possible and has certainly extended the state of the art in proxy credential usage and monitoring. This work will be followed up in numerous ways both directly and indirectly. In direct terms, this work has provided a perfect test bench for research into proxy credential use and potential abuse. A PhD student (Nurazian M Dahalan) is currently undertaking working in this space and has

the completed the first year of her studies. The work has been shown to and accepted by the NGS as adding value to their supporting infrastructure.

Indirectly, the work has solidified the technical work of the teams involved in security. This work is shaping numerous security-oriented projects at NeSC in Glasgow for example where fine grained access control is essential. This expertise in all aspects of security (technical, pragmatic and holistic) in collaborative research environments has shaped many efforts and projects, and is one of the reasons that NeSC in Glasgow continues to thrive with several newly started EU and JISC projects where underpinning security is essential. These include language and literature projects (working with organisations such as the Oxford English Dictionary) through to clinical and biomedical projects involving teams across Europe, e.g. DiPAR (<http://www.dipar.org>) and EuroWABB (<http://www.euro-wabb.org>).

7 References

- D. Reid, R.O. Sinnott, C. Millar, G. Roy, S. Roy, Gordon Stewart, Graeme Stewart, A. Asenov, *Enabling Cutting edge Semiconductor Simulation through Grid Technology*, Journal of the Philosophical Transactions of the Royal Society A, July 2009, 367:2573-2584.
- L Tan, P. Lambert, K. J. Turner, J. Blum, A. Bowes, D. Bell, V. Gayle, S. B. Jones, M. Maxwell, R.O. Sinnott, G. Warner, *Enabling Quantitative Data Analysis through e-Infrastructures*, Social Science Computer Review, January 2009.
- M. Birkin, R. Allan, S. Beckhofer, I. Buchan, J. Finch, C. Goble, A. Hudson-Smith, P. Lambert, R. Procter, D. de Roure, R.O. Sinnott, *The Elements of a Computational Infrastructure for Social Simulation*, UK e-Science All Hands Meeting, Oxford, UK, December 2009 (shortlisted for Journal of the Philosophical Transactions of the Royal Society A).
- R.O. Sinnott, M. Sarwar, A. Kalyanaraman, J.G. Anderson, M. Alexander, J. Greene, *Supporting the Language and Literature Research Community through e-Infrastructures*, UK e-Science All Hands Meeting, Oxford, UK, December 2009.
- A.J. Stell, R.O. Sinnott, J. Jiang, R. Donald, I. Chambers, G. Citerio, P. Enblad, B. Gregson, T. Howells, K. Kiening, P. Nilsson, A. Ragauskas, J. Sahuquillo, I. Piper, *Federating Distributed Clinical Data for the Prediction of Adverse Hypotensive Events* Journal of the Philosophical Transactions of the Royal Society A, July 2009, 367:2679-2690.
- W. Jie, J. Arshad, R.O. Sinnott, *A Review of Grid Authentication and Authorization Technologies and Support for Federated Access Control*, to appear in ACM Computing Surveys, Vol.43, No. 2, January 2011.
- R.O. Sinnott, T. Doherty, D. Martin, C. Millar, G. Stewart, J. Watt, *Supporting Security-oriented Collaborative nanoCMOS Electronics e-Research*, International Conference on Computational Science, Krakow, Poland, June 2008.
- R.O. Sinnott, D. Chadwick, T. Doherty, D. Martin, A. Stell, G. Stewart, L. Su, J. Watt, *Advanced Security for Virtual Organizations: Exploring the Pros and Cons of Centralized vs Decentralized Security Models*, 8th IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2008), May 2008, Lyon, France.
- J. Watt, R.O. Sinnott, T. Doherty, J. Jiang, *Portal-based Access to Advanced Security Infrastructures*, UK e-Science All Hands Meeting conference, Edinburgh, September 2008.
- R.O. Sinnott, T. Doherty, C. Higgins, P. Lambert, S. McCafferty, A. Stell, K. J. Turner, J.P. Watt, *Supporting Security-oriented, Inter-disciplinary Research: Crossing the Social, Clinical and Geospatial Domains*, in Proceedings of International Conference on e-Social Science, Cologne, Germany, June 2009.
- R.O. Sinnott, S. Hussain, *Architectural Design Patterns for Security-oriented Workflows in the Social Science Domain*, Proceedings of International Conference on e-Social Science, Cologne, Germany, June 2009.
- C. Kunz, C. Szongott, J. Wiebelitz, C. Grimm, *Design and Implementation of a Grid Proxy Auditing Infrastructure*, the 5th IEEE International Conference on eScience, Oxford, UK, Dec 2009.
- Globus Developers, Globus Toolkit 5.0.3 source code. source-trees/gsi/gssapi/source/library/globus_i_gsi_gss_utils.c line 521

Project Identifier:
Version: 1.2
Contact: Wei Jie and Richard Sinnott
Date: May 2011

8 Appendices (optional)

A conference paper generated from the project is appended where more technical aspects of the project can be found.

Christopher Bayliss, Richard Sinnott, Wei Jie and Junaid Arshad. The Design, Development and Application of a Proxy Credential Auditing Infrastructure for Collaborative Research, a research paper published on the 5th International Multidisciplinary Conference on e-Technologies (MCETECH 2011), Switzerland, January 2011.