OSG PKI Transition: Experiences and Lessons Learned

Von Welch, Alain Deximo, Soichi Hayashi, Viplav D. Khadke, Rohan Mathure, Robert Quick
Indiana University, USA
E-mail: vwelch@iu.edu, adeximo@iu.edu, hayashis@iu.edu, vdkhadke@indiana.edu, rmathure@umail.iu.edu, rquick@iu.edu

Mine Altunay, Chander S. Sehgal, Anthony Tiradani
Fermi National Accelerator Laboratory, USA
E-mail: maltunay@fnal.gov, cssehgal@fnal.gov, tiradani@fnal.gov

Jim Basney
National Center for Supercomputing Applications, University of Illinois, USA
E-mail: jbasney@illinois.edu

Over the course of 2012-13 the Open Science Grid (OSG) transitioned the identity management system for its science user community from the DOE Grids public key infrastructure (PKI) to a new OSG PKI. This transition was significant in its scope, touching on nearly all aspects of the OSG infrastructure and community. The transition also entailed the adoption of a commercial certificate service as a key component of OSG's PKI.

This transition offers a rare opportunity to better understand identity management and how to prepare for and implement changes in an identity management system. In this paper, we describe OSG's transition and lessons learned from it. We discuss the overall project management approach, including a division of the project into planning, piloting, design, development, implementation and transition phases. We discuss the considered alternatives, both for implementations of the OSG PKI as well as alternatives to a PKI such as federated identity, as well as the criteria we used to make our decision. We conclude with a set of lessons learned from both implementation and in retrospect, and a set of recommendations for other identity systems.

International Symposium on Grids and Clouds (ISGC) 2014
Academia Sinica, Taipei, Taiwan
23-28 March, 2014

1. Introduction

The Open Science Grid[2] (OSG) operates an identity management (IdM) system to allow for authentication of users and services, and to allow for the expression of virtual organization (VO) membership. A key component of the OSG's identity management system is a public key infrastructure (PKI), which generates certificates for users and services that are vetted by a set of trusted agents. (As described subsequently in Section 2, certificates allow users and service to prove their identity to other entities on the OSG.)

From 2003 until 2013, the OSG utilized a PKI operated by the Department of Energy's Energy Science Network (ESnet): the DOE Grids PKI[1][2]. In 2011, ESnet announced it would be shutting down the DOE Grids PKI[3], and ESnet and OSG proceeded to work in collaboration to establish a replacement PKI in the OSG suite of services. This effort was called the OSG PKI Transition Project and ran from November of 2011 through March of 2013. It involved establishing a replacement OSG PKI built around a commercial certificate authority (CA) offering from DigiCert[4]. As of March 23rd, 2013 the DOE Grids PKI stopped issuing new certificates and the OSG PKI is now successfully supporting the identity management needs of the OSG community.

This paper covers the transition process and discusses lessons learned. Lessons are presented to provide guidance both to others who may be undertaking a similar transition as well as those architecting an IdM system that may undertake such a transition in the future. While the OSG PKI Transition involved migration from one PKI to another PKI, the lessons learned are, to varying degrees, applicable to any transition of identity management system.

This paper is organized as follows. We first provide the reader unfamiliar with PKIs a brief background sufficient to understand this paper. Then we describe the DOE Grids PKI and OSG PKI architectures generally, highlighting the changes made by the transition. We describe the transition process with technical, project management and communication challenges. We conclude with lessons learned and a brief discussion of OSG's future plans for IdM.

2. Background: Public Key Infrastructures

In this section we provide readers unfamiliar with PKIs with a brief background sufficient to follow this paper. Note that a PKI is a complicated cryptographic and policy system, and the provided description is both greatly simplified and focused on the context of this paper. For more detail about PKIs please see the Wikipedia article on PKI[5] or [3].

A PKI is a form of identity management system. As an identity management system, its goal is to allow users and systems to authenticate (assert and verify identity) over the Internet. This is accomplished by providing these entities a *certificate*. This certificate is used in a cryptographic

2 https://www.opensciencegrid.org/
3 http://www.es.net/services/grids-service-transition/#Community-comm
4 http://www.digicert-grid.com/
5 http://en.wikipedia.org/wiki/Public-key_infrastructure

manner over a network connection to prove (to some degree of certainty) that the bearer of the certificates holds a certain identity. Other parties use that identity for various functions such as authorization (i.e., determining the rights of the certificate bearer).

An important aspect of any PKI is the process by which entities obtain certificates. Certificates are issued by an entity called a *certificate authority*[6] (CA). CAs are critical in PKIs in that entities trust them to identify and name entities. This trust is a primary factor differentiating PKIs from identity systems based on passwords or cryptographically similar SSH keys[7]. It allows CAs to identify and issue certificates to entities and have those certificates trusted by all members of the community that trust the CA. More technically, it allows a CA to broker trust between entities who do not otherwise know each other.

A challenge CAs have is ensuring that an entity obtaining a certificate legitimately reflects the identity in the certificate, a process called *vetting*. This challenge is accentuated if an entity loses a certificate and needs to be re-issued a certificate with the same identity because that identity typically has been used to establish trust relationships and hence has additional value. Vetting is typically delegated by the CA to *registration authorities*, which represent communities served by the PKI, and trained individuals in those organizations called *registration agents*. In the case of the DOE Grids and OSG PKI, these agents typically have personal knowledge of the entity receiving the certificate and hence are well positioned to identify them.

PKIs maintain a policy describing how they operate (e.g., security controls, vetting processes). For example, each CA issues certificates in a given *namespace* such that no two CAs can issue the same identity, which prevents accidental reuse of identities. One standard set of policies is maintained by the International Grid Trust Federation[8] (IGTF). IGTF policies are in common use in OSG and its collaborators, and form the basis of agreements such that certificates issued under IGTF by different PKIs can be used for interoperability between projects spanning the world and allow for international collaboration, a key to many modern scientific projects.

3. Overview of the OSG PKI and Transition Changes

Figure 1 shows a simplified model of the DOE Grids PKI as it existed from 2003-2013. This PKI was responsible for the issuance and maintenance of over 2000 certificates for users and 9000 certificates for services in over 130 second-level DNS domains (e.g. “university.edu”) within OSG. The components, from bottom of the figure to the top, are:

- A root certificate authority (CA) along with a number of subordinates. The subordinate of most relevance is the DOE Grids CA, which served the OSG community among others. Other subordinate CAs existed for other communities, specialized for their needs, and are outside the scope of this paper.

[6] More accurately *certification authority* but the term *certificate authority* is more commonly used.
[7] http://en.wikipedia.org/wiki/Secure_Shell
[8] http://www.igtf.net/

- A web interface to the DOE Grids CA based on the Red Hat certificate system[9] used by the entities to request, receive and manage their certificate and by Registration Authorities and their Agents to approve requests.
- A set of Registration Authorities for the DOE Grids PKI, of which the Open Science Grid was one. Each authority represents a different community with its own internal processes for approving and issuing certificates.
- The Open Science Grid virtual organizations, each representing a Registration Agent capable of approving certificate issuances for its members.
- The certificates support a set of scientific workflows (e.g. high-through computing jobs, data access) for the different OSG VOs.

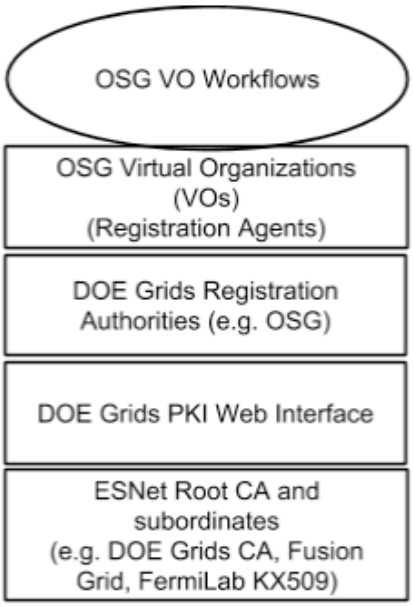

Figure 1: Simplified Model of the DOE Grids PKI which served OSG from 2003-2013.

Figure 2 shows a simplified model of the OSG PKI after the transition in 2013 (for full details on the OSG PKI design, please see [4]). The major changes are:

- The ESnet and DOE Grids CAs have been replaced by a CA service contracted from DigiCert.
- The DOE Grids PKI web interface has been replaced by a web interface integrated into the existing OSG Information Management (OIM) system[10].
- Virtual organizations that were Registration Agents under the DOE Grids-based PKI are rebranded as Registration Authorities.

A requirement of the transition was to not require any modification to the science workflows of the VOs.

[9] http://www.redhat.com/certificate_system/
[10] https://oim.grid.iu.edu/

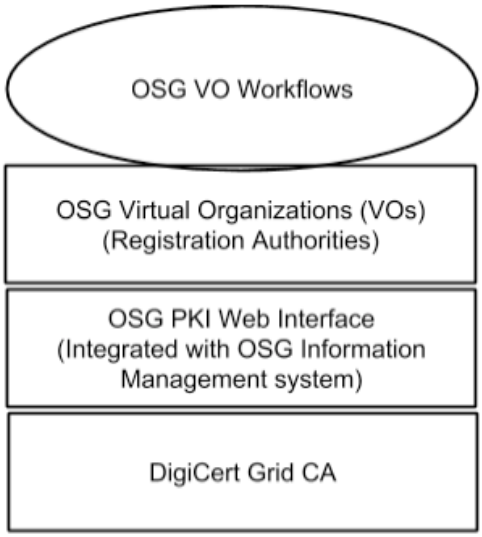

Figure 2: Simplified model of the post-2013 OSG PKI.

A number of organizations which were peers of OSG as clients of the DOE Grids PKI joined the OSG PKI as virtual organizations and Registration Authorities: Argonne National Laboratory, the Earth Systems Grid Federation, the National Fusion Collaboratory, NERSC, and the Oak Ridge National Laboratory.

4. Overview of Transition Project

The OSG PKI Transition Project took approximately 16 months to complete (from November 2011 to March 2013) and had to be undertaken without service interruption to ongoing science being performed by OSG's 50+ virtual organizations. This section presents a brief chronological description of the transition process, highlighting aspects most relevant to the transcition of an identity management system. For full details of the transition project, readers are referred to the project reports[7][8][9][10][11].

4.1 Considered Alternatives and Selection Process

Before the project was initiated, an analysis was undertaken of options available to OSG to replace the DOE Grids CA [5]. A summary of considered options follows:

- A major technological shift away from a PKI was considered (to, e.g., federated identity), and, while viable, was not considered feasible without more planning than could be accomplished before the DOE Grids PKI ceased operations.
- Using non-traditional PKI alternatives (e.g. CILogon [6]) were also considered, but, as discussed in Section 2, IGTF accreditation is important to OSG due to collaborations

with other major infrastructures, namely the Worldwide LHC Computing Grid[11], and again such a transition would take more planning than time allowed. IGTF-accredited options (e.g. CILogon Silver) did not have sufficient maturity.

- OSG assuming operations of the existing DOE Grids PKI was considered, but would have been complicated because it would have to have been untangled from a large ESnet PKI system and the software used for the system was also nearing end of life.
- Not running an OSG PKI and directing the OSG communities to other existing CAs (e.g. the FermiLab KCA[12] and the CERN CA[13]) was considered, but no combination of other CAs, due to their respective policy restrictions, would have provided coverage to the entire OSG community.

While OSG continues to consider and pursue the described alternatives, as described in Section 6, this analysis resulted in the conclusion that OSG needed to establish its own PKI to replace the DOE Grids PKI.

Two major variants of the OSG PKI were considered: OSG setting up and establishing its own CA, or contracting the operation of a CA service from a commercial provider. Establishing and operating an IGTF-accredited CA is a challenge due to the expertise involved in establishing the CA (we refer the reader to [5] for details). Since DigiCert was in progress of obtaining IGTF accreditation for its CA offering, and it was estimated that the cost of contracting with DigiCert was less than the cost of OSG deploying and operating its own CA (purchasing a primary and backup hardware security module, maintaining that module, training and allocating staff to operate it, developing a policy for the CA and getting it accredited by the IGTF, etc.), OSG opted to explore, as described subsequently, the DigiCert commercial CA offering.

4.2 Project Phases

After selecting to explore the DigiCert offering, OSG undertook a three-month Pilot Phase with DigiCert to explore using their CA and ensure it fit the OSG PKI's needs [7]. The Pilot Phase was deemed a success, and OSG undertook the OSG PKI Transition Project to establish a new OSG PKI based on a DigiCert CA and migrate its user community over to the new PKI.

The Transition Project was broken up into discrete phases (Planning Phase, Development Phase, Deployment Phase, Transition, Operations) with a report to OSG management at the end of each phase that served to gate progress [8][9][10]. This separation of the project into phases worked well, with the exception that the Deployment Phase, at one month long, was too short in practice given reporting overhead and was merged with another phase during the project. A final report [11] was produced describing the whole project.

4.3 Major Project Tasks

In this section we present the high level tasks that comprised the OSG PKI Transition project and a brief description of each.

[11] http://wlcg.web.cern.ch/
[12] https://fermi.service-now.com/kb_view.do?sysparm_article=KB0010773
[13] https://gridca.cern.ch/gridca/

Contracting CA services from DigiCert: OSG and DigiCert had to agree on both policies for operations (in order to allow DigiCert to comply with their obligations under the IGTF policies) and a contract for service. Since OSG is not a legal entity itself, the contract was signed on OSG's behalf by one of the member institutions (Indiana University).

Developing a new interface for the PKI: While DigiCert provides a web interface to their PKI (the Managed PKI interface[14]) OSG chose to develop their own interface to a REST API[15] provided by DigiCert. This was done both to allow OSG to tailor the interface to its own needs (particularly command-line clients as described subsequently), integrate it into existing workflows around its VOs and ticketing system, and to allow for supporting other IdM technologies at a future date.

Developing new command-line clients for the PKI: A subset of the OSG community is used to command-line clients (as opposed web browsers) to perform PKI activities. In particular, one use case is *bulk certificate requests*: requests for up to fifty host certificates at a time in order to get certificates for compute clusters comprised of many systems. Development of a new set of command-line tools to replace those developed for the DOE Grids PKI was part of the transition project.

Coordinating the change across the OSG community: Despite the efforts of the OSG project to minimize the work that fell to the community it served, that community still had to undertake numerous changes to adapt to the new PKI, including technical changes (modifying configuration for the new PKI), testing the new OSG PKI certificates with their software stacks, making documentation changes, training their staff for the new PKI interface, and helping their users through the transition. OSG provided guidance to their community on these issues along with communication to keep them abreast of changes.

4.4 Total Project Effort

Total effort on the project was approximately two person-years. This included only the effort for the tasks described in the previous section and does not include the effort put forth by the virtual organizations comprising the OSG community.

A lesson learned was that additional effort would have been beneficial in order to: document all the ways certificates were used in the OSG, developing tests for the new PKI to ensure all aspects of usage were covered, developing documentation for the new PKI, communicating with the OSG community, and helping with the process of developing policies and a contract with DigiCert.

4.5 Contingency Planning

With the change to a new CA provider, the OSG undertook a contingency analysis regarding the risks [12]. This analysis identified a number of risks that could adversely impact the OSG community, including temporary or permanent loss of CA service, a security compromise of the

[14] http://www.digicert.com/ManagedPKI
[15] http://en.wikipedia.org/wiki/Representational_state_transfer

CA, temporary or permanent loss of the web interface, and a security compromise of the web interface. For each, an analysis was done of the impact and a number of options were explored for recovery. A contingency plan [13] was drafted to cover these risks.

5. Lessons Learned from the Transition

In this section we discuss a number of lessons learned from the PKI transition. These lessons are applicable both to projects that may be undertaking similar transitions (between PKIs or from PKI to another form of identity management such as federated identity), or for projects architecting identity management systems that wish to plan for a possible transition in the future. The lessons are listed in no particular order.

5.1 Support Multiple User Identities

Each Certificate Authority has a *namespace* that constrains identities it issues. Namespaces between two different CAs do not overlap to prevent accidental issuance of duplicate identities. This meant that the transition from the DOE Grids PKI to the OSG PKI required a change of identities for all entities served by the OSG.

Applications that allowed binding multiple identities as equivalent (e.g. VOMS[16]) allowed for an easier transition, as identities from both PKIs could both be bound at the same time, allowing entities to transition from an old to new certificate at their leisure. However, applications that only supported one identity per entity (e.g. GUMS[17]) required the change of identity configured in the application to be coordinated with the obtaining of a new certificate, imposing additional overhead and communication, and introducing greater opportunity for problems.

Lesson: Applications should support the ability for entities to have multiple equivalent identities. Those that don't require extra effort for coordination with a change of identity management systems

5.2 Understand Use Cases, Test Early and Often

Testing of software applications that use certificates is important because validation of certificates is a complicated software task. While that task is defined by an IETF Standard (RFC 5280 [14]), unfortunately different applications have different interpretations of the standard that can result in problems when used with different PKIs. The Transition Project attempted to identify all of OSGs use cases involving certificates, however, given the DOE Grids PKI was around for a decade, the diversity and autonomy of the virtual organizations making up the OSG community [15], and the lack of constraints on how user communities could interact with the PKI (e.g. while the interface was technically for web browsers, scripts that interacted directly with the web API were written), some use cases were missed. Late in the Transition Project it was discovered that an important software component (GridSite[18]) demonstrated errors with the new OSG PKI certificates, resulting in urgent software patching.

[16] http://www.italiangrid.it/middleware/voms
[17] https://www.racf.bnl.gov/Facility/GUMS/1.4/
[18] http://www.gridsite.org/

Lesson: Identify all use cases for certificates and test them early with new certificates.

5.3 Early Coordination of Community Critical

Early in the project, the effort required from the community to effect the transition was underestimated. This effort included technical changes (modifying configuration for the new PKI), testing the new OSG PKI certificates with their software stacks, making documentation changes, training their staff for the new PKI interface, and helping their users through the transition.

Lesson: Effort should have been put into place earlier in the project to determine the tasks that the community was responsible for undertaking, communicating that responsibility and providing guidance in its execution.

5.4 Bulk Certificate Request is a Difficult Case

As described previously, OSG supports a use case of requesting multiple certificates at once for clusters and other situations where many machines need certificates. This use case is not one that is (to the best knowledge of the authors) known outside of computational grids such as OSG. The rarity of this use case made it difficult to implement because it was both unfamiliar to DigiCert and not something their interfaces supported naturally (i.e. in an atomic manner), and also because it had a number of difficulties to handle failure modes (e.g. what happens when an error occurs partially through servicing a request).

Lesson: The use case of requesting multiple certificates should have been recognized as being unique and difficult, and more time should have been allocated to plan out its implementation.

5.5 Avoid Browser PKI Functionality

A technical detail not described to this point is that web browsers have the ability to perform operations to assist users in obtaining certificates from PKI web interfaces. This functionality was leveraged in the DOE Grids PKI; however, experience showed that using this functionality limits the PKI's user interface to the specific implementation of the web browser and also introduces problems caused by variances in implementation between the different web browsers (e.g. Internet Explorer, Chrome, Firefox). Given these factors, OSG decided to depart from the use of in-browser PKI functionality and implement the functionality all in the web application. This decision turned out well, reducing overhead in the web application from handling differences in web browser implementations.

Lesson: Not using the in-browser PKI functionality and instead implementing all the PKI functionality in the web application works well.

5.6 In the OSG, Everything is a Virtual Organization

In the original design of the OSG PKI, sites (e.g. DNS domains such as iu.edu, ornl.gov) were treated differently than virtual organizations (e.g. ATLAS, CMS). The plan was that an individual or coherent group would manage each site. In practice this approach turned out to be flawed. Many sites have disjoint groups belonging to multiple virtual organizations and it turned

out identifying the virtual organization involved in a request was critical to route the request to an appropriate agent for vetting. The design was modified during the Transition Project so that sites, in terms of their role in the PKI, were managed by one or more virtual organizations.

Lesson: Certainly in OSG, and probably in most computational grid contexts, virtual organizations are the most important administrative domains, even more so than the organizations administering sites.

5.7 Separation of Web and Grid Certificates

In early discussions with DigiCert, OSG attempted to contract with them as a source of both certificates for Grid services, which as previously described are governed by the IGTF policies, and certificates for secure web browsing (a.k.a. HTTPS or SSL certificates). These latter certificates are governed by more strenuous policies of the CA and Browser (CAB) forum[19]. Limiting the arrangement with DigiCert solely to IGTF certificates simplified both the arrangement and eased OSG's processes for issuing certificates.

Lesson: Separating concerns between Grid and secure web certificates simplified both OSG's relationship with DigiCert and eased OSG's operational procedures.

5.8 Use of A Commercial CA is Practical

Overall the use of a commercial CA was successful in the OSG, but had some particular challenges different from if OSG had established and operated it own CA. These included discussions with DigiCert being slightly complicated because of different nomenclature in use in the two communities, and contract negotiations being complicated because of there being in effect three stakeholders involved (DigiCert, OSG and Indiana University signing on behalf of OSG).

Lesson: Use of a contracted commercial CA worked well, comes with new challenges involved in establishing a business relationship between to different communities.

6. Future OSG Plans for Identity Management

The new OSG PKI has been in operation since March of 2013 and the transition has proven a success. While an IGTF-compliant PKI has served the OSG well now for over a decade, it represents a particular tradeoff of security and usability, driven largely by the need of a subset of OSG's community to be compatible with other projects relying on IGTF. Other infrastructure recently deployed in OSG (e.g. CILogon [6], OSG Connect[20] and GLOW[21]) has demonstrated the demand and practicality of other identity management systems that offer improved usability. Hence OSG continues to explore these alternatives, and others, including usernames and passwords, to best meet the needs of its user community.

[19] https://cabforum.org/
[20] http://osgconnect.net/
[21] http://research.cs.wisc.edu/htcondor/glow/

7. Acknowledgments

We thank the Department of Energy, the ATLAS and CMS projects, the Indiana University Center for Applied Cybersecurity Research, and ESnet for supporting the work described in this paper.

Additionally we thank Tim Cartwright, Jeremy Fischer, John Hover, Christiane A. Ludescher-Furth, Dhiva Muruganantham, Ruth Pordes, Alain Roy, Lauren Rotman, Mátyás Selmeci, and John Volmer for contributions to this work. We also thank all the OSG virtual organizations who worked to make the transition a success for their user communities.

References

[1] T. Genovese, DOE's PKI Service for Grids. Terena Second Authentication and Authorisation Workshop. Malaga, Spain. 20-21 November 2003. http://www.terena.org/activities/tf-aace/AAworkshop/ppt/US-DOE-PKI_TG%20.ppt

[2] DOEGrids Certificate Service. http://www.doegrids.org. Visited March 16, 2014.

[3] Adams, Carlisle & Lloyd, Steve (2003). Understanding PKI: concepts, standards, and deployment considerations. Addison-Wesley Professional. pp. 11–15. ISBN 978-0-672-32391-1

[4] OSG PKI Design. https://confluence.grid.iu.edu/display/CENTRAL/Design+Document

[5] OSG ID Mgmt CA Replacement Plans. http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1077

[6] Jim Basney, Terry Fleury, and Jeff Gaynor, "CILogon: A Federated X.509 Certification Authority for CyberInfrastructure Logon," XSEDE Conference, July 2013, San Diego, CA. http://dx.doi.org/10.1145/2484762.2484791

[7] Mine Altunay, Jim Basney, Jeremy Fischer, Chander Sehgal and Von Welch. *OSG DigiCert Pilot Report*. OSG-doc-1097, March2012. http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1097

[8] OSG PKI Planning Phase Report. http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1120

[9] OSG PKI Development and Deployment Phase Report. http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1145

[10] OSG PKI Transition Phase Report. http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1148

[11] OSG PKI Transition Final Report. http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1156

[12] Mine Altunary and Von Welch. OSG PKI Contingency Analysis. OSG-doc-1115, July 2012. http://www.vonwelch.com/pubs/OSGPKIContingencyAnalysis12

[13] Mine Altunary and Von Welch. OSG PKI Contingencies-- Recovery Plan. OSG-doc-1121, July 2012. http://www.vonwelch.com/pubs/OSGPKIContingencyPlan12

[14] D. Cooper, et al. RFC 5280: Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. IETF, May 2008.

[15] OSG Document 18-v12: A Blueprint for the Open Science Grid. March 2011. http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=18