# International Grid CA Interworking, Peer Review and Policy Management through the European DataGrid Certification Authority Coordination Group

J. Astalos[13], R. Cecchini[14], B. Coghlan[6], R. Cowles[20], U. Epting[11], T. Genovese[8], J. Gomes[15], D. Groep[18], M. Gug[9], A. Hanushevsky[20], M. Helm[8], J. Jensen[3], C. Kanellopoulos[1], D. Kelsey[3,⋆], R. Marco[12], I. Neilson[9], S. Nicoud[5], D. O'Callaghan[6], D. Quesnel[2], I. Schaeffner[11], L. Shamardin[16], D. Skow[10], M. Sova[4], A. Wäänänen[17], P. Wolniewicz[19], and W. Xing[7]

[1] Aristotle University of Thessaloniki, Greece
[2] Canarie, Canada
[3] Rutherford Appleton Laboratory, UK
[4] CESNET, Czech Republic
[5] CNRS/UREC CPPM, France
[6] Trinity College Dublin, Ireland
[7] University of Cyprus, Cyprus
[8] ESnet/LBNL, USA
[9] European Organization for Nuclear Research (CERN), Switzerland
[10] Fermi National Accelerator Laboratory, USA
[11] Forschungszentrum Karlsruhe, Germany
[12] Instituto de Física de Cantabria (CSIC-UC), Spain
[13] Slovak Academy of Sciences, Slovakia
[14] INFN, Italy
[15] Laboratório de Instrumentação e Física Experimental de Partículas, Portugal
[16] Moscow State University, Russia
[17] Niels Bohr Institute, Denmark
[18] NIKHEF, Netherlands
[19] Poznań Supercomputing and Networking Center, Poland
[20] Stanford Linear Accelerator Center, USA

**Abstract.** The Certification Authority Coordination Group in the European DataGrid project has created a large-scale Public Key Infrastructure and the policies and procedures to operate it successfully. The infrastructure demonstrates interoperability of multiple certification authorities (CAs) in a novel system of peer-assessment of the roots of trust. Crucial to the assessment is the definition of minimum requirements that all CAs must meet in order to be accepted. The evaluation is aided by software-generated trust matrices. Related work building on this infrastructure is described. The group's policies and experience now form the basis of the new European Policy Management Authority for Grid Authentication in e-Science.

⋆ Corresponding author: D.P.Kelsey@rl.ac.uk

## 1 Introduction

This paper describes the creation and operation of a Public Key Infrastructure (PKI) for grid authentication used by several international grids. The European DataGrid (EDG) project[1], which began in January 2001, was the first European project to establish a wide-scale grid. During the three years of EDG, the authentication requirements of this and other grid projects led to the inclusion of 21 Certification Authorities (CAs) in the PKI. These CAs provide authentication services for people and grid services in the majority of the EU member states and also in Canada, Russia, Taiwan and the USA.

EDG was the first grid project to involve more than a small number of nations, each with their own administrative and security domains. Initially this was not perceived as an issue, but project members soon realised that the resource owners required a more structured approach to security. The Certification Authority Coordination Group (CACG) was established at the beginning of the project to define a common authentication infrastructure trusted by all relying parties that were part of the EDG project. EDG's sister projects, such as DataTAG[2] and CrossGrid[3] have adopted the EDG security model. GridLab[4] recognises the CACG member CAs. The LCG[5] and EGEE[6] projects also take the EDG approach to authentication.

EDG security activities fell into three categories: authentication, authorization and coordination. EDG decided to keep authentication and authorization separate (due to the more dynamic nature of authorization) while recognizing that authentication often includes some implicit authorization. Authentication was based on the Globus Grid Security Infrastructure (GSI)[7]. The Security Coordination Group (SCG) have documented the EDG authorization developments [8,9,10].

## 2 DataGrid Authentication

The EDG SCG collected and documented the security requirements of the project [11]. These included 17 requirements for authentication of which three important items were: for a user to authenticate just once per session; interoperable authentication between many grids and applications; and the ability of authentication to be revoked in the event of loss or compromise of an identity credential. The requirements led to the use of Globus GSI. This uses a Public Key Infrastructure (PKI) with X.509 certificates[12]. Identity is checked by a Registration Authority (RA) and certified by a Certification Authority (CA). Users, hosts and services perform mutual authentication. Delegation with short-lifetime proxy credentials achieves the important goal of single sign-on[13]. A grid mapfile maps a certificate's distinguished name (DN) to a local account and authorization is enforced by the local security mechanisms.

The CA Coordination Group had the task of creating a PKI, which was unique in its successful coordinated use of the technology with a large number of independently operated CAs. The infrastructure was for GSI authentication only:

it specifically did not support long-term encryption or digital signatures. A single certification authority for the whole project was not thought to be sufficient due to concerns about a single point of failure or attack. It was also important to have robust relationships between each CA and its associated RAs. To meet these requirements, an appropriate scale was one CA for each country, large region or international organization. A single hierarchy would have excluded some pre-existing CAs, reduced the ability of CAs to meet local needs, and was not convenient to support with the Globus software. For these reasons a coordinated group of peer CAs was the most suitable choice. The EDG project did not have any resources allocated to run such a PKI, so efforts were drawn from participating national projects and organizations.

### 2.1 Globus Grid Security Infrastructure Features

In Globus GSI the end-entity certificate is used to sign a 'proxy' certificate. In the validation of the proxy certificate, the end-entity basicConstraints (which state the that certificate is not a CA certificate) are deliberately ignored, a violation of the normal validation procedures. GSI proxy certificates are now an IETF standard described in RFC 3820 [14].

X.509 CRLs[15] have a nextUpdate field that conveys a hint when a new CRL can be obtained. In GSI, this field is interpreted strictly as an expiration date: if the CRL for a particular CA is present but outdated, end-entity certificates signed by this CA will not be accepted by the software.

### 2.2 Status of DataGrid PKI

At the end of the EDG project there were 21 approved national certification authorities. CNRS, France ran a 'catch-all' CA, for those without a national CA, with appropriate RA mechanisms. In Table 1, 'Total Issued' certificates include those for users, hosts and services and also includes certificates which have since expired or been revoked. In the 'Currently Valid' column is the current number of active certificates. The data for this table were collected in April 2004. Note: The CERN CA serves the CERN community. The FNAL Root CA is accredited but only issues CA certificates.

The CAs in the PKI each provide an equivalent service but with different resources. Many CAs use OpenSSL[16] . Others use Globus Simple CA[17] and various versions of OpenCA[18]. The DOEGrids CA uses Sun ONE[19] Certificate Server.

The relying parties, i.e. users, services and resources, of the PKI must be able to download and install the CA certificates, namespace signing policies, and CRLs of each trusted CA in a secure and robust way. The Certification Authority Repository[21] provides this information for grid administrators. CA information is distributed in RPM (RedHat Package Manager) format for the EDG testbed. Scripts have been written to update CRLs periodically, as they are not fetched automatically by the Globus software.

[21] Certification Authority Repository: http://marianne.in2p3.fr/datagrid/ca/

**Table 1.** certification authority statistics

| CA | Country | Total Issued | Currently Valid |
|---|---|---|---|
| ArmeSFo | Armenia | 1 | 1 |
| ASCCG | Taiwan | 80 | 68 |
| CERN | CERN | 640 | 321 |
| CESNET | Czech Republic | 365 | 211 |
| CNRS | France & Catch-all | 1400 | 392 |
| CyGrid | Cyprus | 18 | 14 |
| DataGrid-ES | Spain | 408 | 191 |
| DOEGrids | USA | 2807 | 1572 |
| GridCanada | Canada | 570 | 467 |
| Grid-Ireland | Ireland | 170 | 111 |
| GridKA | Germany | 364 | 225 |
| HellasGrid | Greece | 49 | 33 |
| INFN | Italy | 1956 | 1158 |
| LIP | Portugal | 61 | 43 |
| NIKHEF | Netherlands | 321 | 124 |
| NorduGrid | Nordic Countries | 579 | 316 |
| PolishGrid | Poland | 266 | 207 |
| Russian DataGrid | Russia | 230 | 99 |
| SlovakGrid | Slovakia | 26 | 18 |
| UK e-Science CA | UK | 1856 | 1297 |
| Total | | 12167 | 6868 |

## 3 Minimum Requirements for Grid Certification Authorities

One of the major activities of the CACG has been the production and maintenance of a set of minimum requirements and best practices for an "acceptable and trustworthy" CA as defined by the relying parties of EDG and related grid projects, taking into account the level of risk associated with the assets to be protected. These requirements have evolved during the project largely as a result of the numerous difficulties that arise when interoperating between different linguistic, administrative, networking and security domains. This section is based on the Minimum Requirements document of the European Policy Management Authority for Grid Authentication in e-Science (EUGridPMA), which is publicly available from http://www.eugridpma.org/.

In this section, the key words 'MUST', 'MUST NOT', 'REQUIRED', 'SHALL', 'SHALL NOT', 'SHOULD', 'SHOULD NOT', 'RECOMMENDED', 'MAY', and 'OPTIONAL' are to be interpreted as described in RFC 2119. Text in *italics* provides discussion and clarification of the requirements.

Due to certain idiosyncrasies of the grid middleware, the PKI structure SHOULD NOT follow the conventional hierarchical model: there SHOULD be one certification authority (CA) per country, large region or international organization each with an associated network of registration authorities (RA). The RAs

handle the tasks of validating the identity of the end entities and authenticating their requests, which will then be forwarded to the CA. The CA will handle the tasks of issuing CRLs; signing certificates and CRLs; and revoking certificates when necessary.

Requirements of the Certification Authority:

**Computer Security Controls:** The CA computer, where certificates are signed, SHOULD be a **dedicated machine**, running only services needed for CA operations. It MUST be located in a secure environment where access is limited to specific trained personnel and MUST be kept disconnected from any kind of network. If the CA computer is equipped with at least a FIPS 140-1 level 3 Hardware Security Module or equivalent it MAY be connected to a highly protected/monitored network. The security controls MUST be documented and the documentation made available to the PMA.

**CA Namespace:** Each CA MUST sign only a well defined namespace that does not clash with any other CA.

**Policy Document & Identification:** Every CA MUST have a Certification Policy and Certification Practice Statement (CP/CPS) and assign it an OID (object identifier). Whenever there is a change in the CP/CPS the OID of the document MUST change and changes MUST be approved by the PMA before signing any certs under the new CP/CPS. All the CP/CPSs under which valid certs are issued MUST be available on the web. *We currently recommend the RFC 2527 template for the CP/CPS document.*

**CA Key:** The CA Key MUST have a minimum length of 2048 bits and, for CAs that issue end-entity certificates, the lifetime MUST be no longer than 5 years and no less than twice the maximum life time of an end-entity certificate. The private key of the CA MUST be protected with a pass phrase of at least 15 characters and known **only** by specific personnel of the certification authority. A copy of the encrypted private key MUST be kept on an offline medium in a secure place. The pass phrase of the encrypted private key MUST also be kept on an offline medium in a secure place, separate from the key.

**CA Certificate:** The CA certificate MUST have the extensions keyUsage and basicConstraints marked as critical.

**CRLs:** The maximum CRL lifetime MUST be at most 30 days and the CA MUST issue a new CRL at least 7 days before expiration and immediately after a revocation. The CRLs MUST be published in a repository accessible via the World Wide Web, as soon as issued. *We recommend that all relying parties update their local copies of CRLs at least once per day.*

**Records Archival:** The CA MUST record and archive all requests for certificates, along with all the issued certificates; all the requests for revocation; all the issued CRLs; and the login/logout/reboot records of the issuing machine.

**Key Changeover:** The CA's private signing key MUST be changed periodically; from that time on only the new key will be used for certificate signing purposes. The overlap of the old and new key MUST be at least the longest time an end-entity cert can be valid. The older but still valid certificate MUST be available to verify old signatures, and the private key to sign CRLs, until all the certificates signed using the associated private key have expired.

**Repository:** The repository MUST be run on a best-effort basis, with an intended availability of 24×7.

**Compliance Audits:** Each CA MUST accept being audited by other CAs to verify its compliance with the rules and procedures specified in its CP/CPS document.

**Operational Audits:** The CA MUST perform operational audits of the CA and RA staff at least once per year.

Requirements of the Registration Authority:

**Entity Identification:** In order for an RA to validate the identity of a person, the subject MUST contact the RA personally and present photographic identification and/or valid official documents showing that the subject is an acceptable end entity as defined in the CA's CP/CPS. In case of host or service certificate requests, the request MUST be delivered to the RA by the person in charge of the specific entities using a secure method.

**Name Uniqueness:** The subject name listed in a certificate MUST be unambiguous and unique for all certificates issued by the CA.

**Records and Archival:** The RAs MUST record and archive all requests and confirmations.

**Communication with CA:** The RA MUST communicate with the CA with secure methods that are clearly defined in the CP/CPS. *e.g. signed emails, voice conversations with a known person, SSL protected mutually authenticated private web pages .*

The end-entity (EE) keys MUST be at least 1024 bits long and MUST NOT be generated by the CA or the RA. The EE certificates MUST have a maximum lifetime of 1 year and MUST NOT be shared among end entities. The EE certificate MUST contain information to identify which CP/CPS was used to issue the certificate (e.g. OID or date). The extensions basicConstraints and keyUsage MUST be marked as critical and the basicConstraints MUST be set to "CA: False". The CA SHOULD make a reasonable effort to make sure that end-entities understand the importance of protecting their private key, with a pass phrase of at least 12 characters.

## 4 Trust Evaluation

To establish trust, each CA is required to demonstrate to the group that the setup and policies are secure. This is usually done in person at a meeting of the CACG where detailed questions about the CP/CPS, the practices, the RA structure, etc. are answered. After satisfying this peer review a CA will be 'accredited'. Each relying party (RP) wants to evaluate all the CAs, either that they meet the RP's standard, or that they meet an agreed common standard. The CACG peer review establishes this common standard. This requires inspection of each CA's CP/CPS by a volunteer subset of the other CAs. Third-party audits have been considered but would be time-consuming and expensive and none have

yet been done. Evaluation of trust is a continuous and long-term process and experience has shown that personal contacts are fundamental. The Global Grid Forum (GGF)[20] has established several working groups to establish policies and procedures in this area.

The assessment process is manual, and CA managers want to make it more automatic. Software is being developed to aid this process based on evaluation of a CA *Feature Matrix*. CP/CPS documents are encoded in a report file and the Feature Matrix displays the features. The CA report file uses a basic contextual language involving key-value pairs, e.g. name = 'CERN CA'. The language is designed to enable later extension to allow formal analysis, but is presently very simple. Features can be evaluated relative to *rulesets*. A *default ruleset* has been defined for EDG, based on the CACG minimum requirements. This allows the construction of a CA *Acceptance Matrix*.[22] The GGF concept of assurance levels is accommodated to allow rulesets to be defined for each level[21]. Each Virtual Organization (VO) can also define their own rules that override and extend the default ruleset. The *Ruleset Inclusion Principle* extends from the general to the specific. It can be extended to CAs, sites, hosts, users and even specific services simply by defining the appropriate ruleset. Thus a typical chain might be: default ruleset → VO ruleset → host ruleset. It is not necessary for a subject to have all possible rulesets in their possession, only those rulesets that they are interested in. Further evaluations with example user, host and service certificates, and samples of issued certificates are possible and this is the current focus. There are other complementary approaches: for example, evaluating an XML encoding of a CP/CPS[22].

## 5 Related Work

### 5.1 Certificate Request Applets

Java applets have been developed to be used for certificate requests. An applet generates the keys and associated request and submits them to the CA. Another applet is used to download the certificate and match it with the corresponding private key. Once the certificate and key have been matched, they are exported in PKCS #12 format which can be imported into a browser.

The applets must be signed, since they read and write files on the user's disk, and so that the user trusts that the applets were issued by the CA. An advantage of using applets is that the CA can perform some basic validation when the user applies for the certificate, rather than rejecting invalid requests at a later stage. The applet method allows the CA to check the strength of the user's passphrase without ever seeing the passphrase or the private key. This is a great advantage over the 'normal' method where the user must be trusted to generate a sufficiently strong passphrase.

[22] CA Trust Matrices: http://www.cs.tcd.ie/coghlan/cps-matrix/cps-matrix.cgi

### 5.2 Compromised and Exposed Private Keys

The CACG has explored the issues related to the compromise and exposure of private keys. Compromised keys should be revoked, but the definition of a 'compromise' of credential confidentiality is unclear. It is *a priori* impossible to prove confidentiality to a third party, so we must rely on best professional judgement. This necessarily means cases will have to be evaluated individually. The following cases provide a working definition for 'compromise' and 'exposure' of private keys.

If a private key can be shown to be in the possession of someone other than the user then it is considered 'compromised'. When an attacker has had access to the user's unencrypted private key, it will be considered a 'compromise' unless forensic analysis can rule out access to the key. Compromised keys must be revoked.

If an encrypted private key is available to someone other than the user then it is considered 'exposed'. An encrypted private key is vulnerable to offline attack, protected only by the user-chosen passphrase. When an attacker has had access to the user's encrypted private key and the attacker demonstrates sufficient skill and knowledge of PKI, it will be considered a 'compromise' unless forensic analysis can rule out access to the key. Exposed keys should be reported to the appropriate CA who will alert the user to the exposure. Keys visible in the course of system administration will not normally be considered exposed.

### 5.3 Online Certificate Services

Traditionally, grid certification authorities have been operated offline. This reduces the risk of compromise of the CA signing key. Online certificate services are those which store private keys, and generate or sign certificates on a network-connected system. LCG[5] is using a KCA and ESnet is proposing a minimum requirements profile for online services.

**5.3.1 Kerberized Certification Authority** The Kerberized Certification Authority (KCA) provides a automated mechanism for an organization with an existing Kerberos infrastructure to generate X.509 credentials for use in PKI-based authentication systems. The KCA software is distributed by the NSF Middleware Initiative (NMI)[23].

The KCA consists of a secure server which communicates with a client to generate PKI credentials. The KCA service is attractive for sites operating a Kerberos-based authentication infrastructure. The user is not issued a long term private key and proxy maintenance uses the existing Kerberos infrastructure. The administrative overhead and possibility of error or deliberate attack on another RA is removed. Since the KCA issues only short-lived certificates, there is no need to distribute CRLs. Compared to a well run offline service the danger of signing key compromise is increased. In the context of long-running jobs in the grid the problem arises of how to renew a proxy certificate derived from a user's Kerberos token which is typically valid for about one day.

**5.3.2 Virtual Smart Cards** The SLAC Virtual Smart Card system[24], provides an online credential store analogous to a physical smart card. As users cannot be trusted to keep private keys secure they should not be given the private key. VSC can provide stronger security guarantees with a central restricted-access server than individual untrustworthy users, and it allows users to generate proxy certificates from anywhere that has access to the VSC server. The disadvantages are that the private keys are concentrated in one place, therefore giving a single point of failure, and the authentication for the whole system is only as strong as the authentication with the VSC server, so this must be of high quality, e.g. a well-administered Kerberos setup.

## 6 Summary

During the last three years the Certification Authorities Coordination Group has successfully built a large-scale Public Key Infrastructure which is now in global production use. This infrastructure allows users and services to have just one identity credential which is accepted and trusted by a growing number of VOs and grid projects.

The evolution of the best practices, minimum requirements and the associated establishment of inter-domain trust via peer review on behalf of the various relying parties, has taken time and involved many debates during the meetings of the group. The tools developed for trust evaluation and the various technical challenges of grid authentication have enabled the group to avoid having to spend all of its time concentrating on policies and procedures. As described in the paper, future work building on this infrastructure has already started. The expected growth of online certificate services and repositories, together with online certificate status checking, is likely to play a significant role in future authentication services.

The policies of the Certification Authority Coordination Group worked extremely well for EDG. With the input from other grid projects it has become a large group and now forms the basis of the new European Policy Management Authority for Grid Authentication in e-Science (EUGridPMA)[25]. This new body, which is initially coordinating authentication services for EGEE[6], DEISA[26], LCG[5] and SEEGRID[27], is associated with the global Grid Policy Management Authority[28] initiative, started in 2002 to coordinate the PMAs. The policies, procedures and technical solutions developed by CACG and described in this paper, are being taken forward by the EUGridPMA with the aim of turning this into an even more pervasive general infrastructure for authentication for e-Science.

The authors wish to acknowledge the EU and many national funding bodies, institutes and projects that allowed their staff to participate in the activities of the Certification Authority Coordination Group. We thank all our colleagues in each of the grid projects, particularly European DataGrid, for providing very valuable comments and feedback on the authentication infrastructure during the project.

## References


1. European DataGrid. (2004) http://www.edg.org/.
2. DataTAG. (2004) http://datatag.web.cern.ch/.
3. CrossGrid. (2004) http://www.crossgrid.org/.
4. GridLab. (2004) http://gridlab.org/.
5. LHC Computing Grid. (2004) http://lcg.web.cern.ch/.
6. Enabling Grids for E-science in Europe. (2004) http://www.eu-egee.org/.
7. Foster, I., Kesselman, C., Tsudik, G., Tuecke, S.: A security architecture for computational grids. In: ACM Conference on Computers and Security. ACM Press (1998) 83–91
8. DataGrid Security Coordination Group: Security Design. (2003) https://edms.cern.ch/document/344562.
9. DataGrid Security Coordination Group: Final Security Report. (2004) https://edms.cern.ch/document/414762.
10. Cornwall, L.A. *et al.*: Security in multi-domain grid environments. Journal of Grid Computing (2004)
11. DataGrid Security Coordination Group: Security Requirements Testbed 1 Security Implementation. (2002) https://edms.cern.ch/document/340234.
12. IETF: PKIX Charter. (2004) http://www.ietf.org/html.charters/pkix-charter.html.
13. Butler, R., Engert, D., Foster, I., Kesselman, C., Tuecke, S., Volmer, J., Welch, V.: Design and deployment of a national-scale authentication infrastructure. IEEE Computer **33** (2000) 60–66
14. Tuecke, S., Welch, V., Engert, D., Pearlman, L., Thompson, M.: Internet X.509 Public Key Infrastructure Proxy Certificate Profile. (2003) http://www.ietf.org/internet-drafts/draft-ietf-pkix-proxy-10.txt.
15. Housley, R., Polk, W., Ford, W., Solo, D.: Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. (2002) RFC 3280.
16. OpenSSL. (2004) http://www.openssl.org/.
17. Globus Simple CA. (2004) http://www.globus.org/security/simple-ca.html.
18. OpenCA. (2004) http://www.openca.org/.
19. Sun Open Network Environment. (2004) http://wwws.sun.com/software/sunone/.
20. Global Grid Forum. (2004) http://www.ggf.org/.
21. Butler, R., Genovese, T.: Global Grid Forum Certificate Policy Model. (2003)
22. Ball, E., Chadwick, D., Basden, A. In: The Implementation of a System for Evaluating Trust in a PKI Environment. Volume 2 of Evolaris. SpringerWein (2003) 263–279
23. NSF Middleware Initiative. (2004) http://www.nsf-middleware.org/.
24. Hanushevsky, A., Cowles, R.: Virtual Smart Card. (2002) http://www.slac.stanford.edu/ abh/vsc/.
25. European Grid Policy Management Authority for e-Science. (2004) http://www.eugridpma.org/.
26. Distributed European Infrastructure for Supercomputing Applications. (2004) http://www.deisa.org/.
27. South Eastern European Grid-enabled eInfrastructure Development. (2004) http://www.see-grid.org/.
28. GridPMA. (2004) http://www.gridpma.org/.