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

J. Astalos[13], R. Cecchini[14], B.A. Coghlan[6], R.D. Cowles[20], U. Epting[11], T.J. Genovese[8], J. Gomes[15], D. Groep[18], M. Gug[9], A.B. Hanushevsky[20], M. Helm[8], J.G. Jensen[3], C. Kanellopoulos[1], D.P. Kelsey[3*], R. Marco[12], I. Neilson[9], S. Nicoud[5], D.W. 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, GR 541 24 Thessaloniki, Greece.*
[2] *Canarie, 110 O'Connor St., 4th floor, Ottawa, Ontario, K1P 5M9, Canada.*
[3] *CCLRC, Rutherford Appleton Laboratory, Chilton, Didcot, OX11 0QX, UK.*
[4] *CESNET z. s. p. o., Zikova 4, Praha 6, 160 00 , Czech Republic.*
[5] *CNRS/UREC CPPM, 163 av de Luminy, case 902 F-13288 Marseille Cedex 09, France.*
[6] *Department of Computer Science, Trinity College, Dublin 2, Ireland.*
[7] *Department of Computer Science, University of Cyprus, 75 Kallipoleos Street, PO Box 20537, CY-1678 Nicosia, Cyprus.*
[8] *ESnet/LBNL, 1 Cyclotron Road, MS 50B-4230 Berkeley, CA 94720, USA.*
[9] *European Organization for Nuclear Research (CERN), CH-1211 Genève 23, Switzerland.*
[10] *Fermi National Accelerator Laboratory, PO Box 500 Batavia, IL 60510-0500, USA.*
[11] *Forschungszentrum Karlsruhe, Postfach 3640, D-76021 Karlsruhe, Germany.*
[12] *Instituto de Física de Cantabria (CSIC-UC), Avda de los Castros s/n 39005, Santander, Spain.*
[13] *Institute of Informatics, Slovak Academy of Sciences, Dubravska cesta 9, 845 07 Bratislava, Slovakia.*
[14] *INFN, Sezione di Firenze, Via G. Sansone 1, I 50019 Sesto Fiorentino, Florence, Italy.*
[15] *Laboratório de Instrumentação e Física Experimental de Partículas, Av. Elias Garcia 14, 1ž 1000-149 Lisbon, Portugal.*
[16] *Skobeltsyn Institute of Nuclear Physics, Moscow State University, Vorobjovy Gory, Moscow, 119992, Russia.*
[17] *NBI, Blegdamsvej 17, DK-2100 Copenhagen, Denmark.*
[18] *NIKHEF, PO Box 41882, NL-1009DB Amsterdam, Netherlands.*
[19] *Poznań Supercomputing and Networking Center ul. Noskowskiego 10 61-704 Poznań, Poland.*
[20] *Stanford Linear Accelerator Center, 2575 Sand Hill Road Menlo Park, CA 94025, USA.*

[*] *corresponding author. Email: D.P.Kelsey@rl.ac.uk*

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

Keywords: Authentication, Certification Authority, Globus, GSI, Public Key Infrastructure, Security, Trust

1. Introduction

This paper describes the creation and successful operation of a unique Public Key Infrastructure (PKI) used for the purposes of grid authentication by several large international grids. The European DataGrid (EDG) project [13], which started in January 2001, was the first European project to establish a wide-scale Grid. During the three years of the EDG project, the authentication requirements of EDG and other related Grid projects led to the inclusion, by the end of 2003, of 21 Certification Authorities (CAs) in this PKI. These include CAs providing authentication services for people and grid services in the majority of the EU member states, in many of the states joining the EU in 2004, and also in Canada, Russia, Taiwan and the USA.

The European DataGrid 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, and was not covered by any task in the work-plan. Project members realised very quickly that the resource owners required a more structured approach to security. The Certification Authority Coordination Group (CACG) was therefore established at the beginning of the project to define a common authentication infrastructure that was trusted by all relying parties that were part of the EDG project. DataGrid has since fostered several sister projects, such as DataTAG[10] and CrossGrid[6], which are comparable in size, and have adopted the DataGrid security model. GridLab [18] also recognises the CACG member CAs. The new larger LCG [23] and EGEE [12] projects also take the DataGrid approach to authentication.

In DataGrid, each middleware work package was responsible for including appropriate security features. DataGrid security activities fell into three categories: authentication, authorisation and coordination. Authentication is based on the Globus Grid Security Infrastructure (GSI) [15] with added DataGrid-specific functionality. DataGrid decided to keep authentication and authorisation separate (due to the more dynamic nature of authorisation) while recognizing that authentication often includes some implicit authorisation. The Security Coordination Group have coordinated and documented the authorisation developments in DataGrid [8, 9, 5].

Many security terms mentioned in this paper are defined in [29].

2. DataGrid Authentication

The EDG Security Coordination Group collected and documented the security requirements of the project[7]. These included 17 requirements for authentication of which three important items were:the need of a user to authenticate just once per session; for interoperable authentication between many Grids and applications; and for the ability of authentication to be revoked in the event of loss or compromise of an identity credential. These requirements resulted in the use of an authentication infrastructure based on the Globus Grid Security Infrastructure (GSI). GSI uses a modified form of Public Key Infrastructure (PKI) with X.509 certificates[22]. Identity is checked by a Registration Authority (RA) and certified by a Certification Authority (CA). Users, hosts and services perform mutual authentication. Delegation is via proxy credentials that have a limited lifetime. This distributed form of authentication achieves the important goal of single sign-on[3]. A grid mapfile maps a certificate's distinguished name (DN) to a local Unix or Kerberos [24] user and authorisation is then enforced by the local security mechanisms.

The CA Coordination Group had the task of creating an actual PKI, which was unique in its successful use of the technology with a large number of independently operated CAs. The infrastructure was to be used for grid authentication only, and then only in the context of distributed resource access through Globus GSI. It specifically did not support long-term encryption of data or digital signatures.

A single certification authority for the whole project would not be sufficient due to concerns about a single point of failure or attack. It was also considered important to achieve robust relationships between the CA and associated RAs. To meet these requirements it was decided that an appropriate scale was one CA for each participating country. Hierarchical or cross-signed arrangements of multiple CAs are not compatible with Globus GSI, so a coordinated group of peer CAs appeared to be 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 organisations.

2.1. Globus Grid Security Infrastructure Features

In Globus GSI, to achieve single sign-on, the end-entity certificate is used to sign a ‘proxy’ certificate. In the validation of this ‘proxy’ certificate, the end-entity basicConstraints (which state the that certificate is not a CA certificate) are deliberately ignored, this being a violation of the normal validation procedures. GSI Proxy certs are currently being standardised in the IETF (Internet Engineering Task Force) PKIX standards group[32].

X.509 certificate revocation lists (CRLs) [21] have a nextUpdate field that normally conveys a hint when a new CRL could be obtained, and generally the CRL should not be used after this date, although the X.509 standard is ambiguous on this point. 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. Current Status of DataGrid PKI

Currently there are 21 approved national certification authorities. Each includes registration authorities that check identity. During the EDG project CNRS (France) is acting as a ‘catch-all’ CA with appropriate RA mechanisms. CA managers check each other against an agreed list of minimum requirements (see Section3). Software is being developed to aid this process (see Section 4).

In Table I, ‘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. In Table II, the root certificate fingerprints in MD5 and SHA1 format are published for the record.

2.3. Certification Authority Software

The CAs in the coordination group PKI each provide an equivalent service but with different resources. The software used by each CA is not imposed by the group. A number of certification authorities use OpenSSL [27] to accept requests, sign certificates and issue revocation lists. The commands are usually wrapped into Bash or Perl scripts for ease of use. The Globus Simple CA [17] software is also in use and various versions of OpenCA [26] are used by a number of CAs. The DOEGrids CA uses Sun ONE [31] Certificate Server.

Table I. Certification Authority Statistics

CA	Country[*]	Total Issued	Currently Valid
ArmeSFo[†]	Armenia	0	0
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
FNAL[†]	USA	1	1
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

[*] CERN is an international organization. ‘Catch-all’ serves those without a national CA.
[†] The ArmeFSo CA has not yet become fully operational. The FNAL Root CA does not issue end-entity certificates, only CA certificates.

2.4. Certification Authority Repository

The various 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 Globus GSI signing policy file must be configured with the namespace and the distinguished name of each CA. The Certification Authority Repository[1] provides this information for grid administrators.

CA information is distributed in RPM (RedHat Package Manager) format to allow easy installation on the EDG testbed. Scripts have been written to update CRLs periodically, as they are not fetched automatically by the Globus software. The repository is a central place for CA contact information, links to documentation, certificates, CRLs and RPMs. One change that should be made in future projects is to make CA package releases independently of the software release schedule, to permit addition or modification of the CA packages as necessary.

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. Acceptable and trustworthy is defined from the point of view of the relying parties of EDG and related grid projects taking into account the level of risk associated with the assets the projects have to protect. These minimum requirements have evolved over the life of the EDG project in an iterative discursive fashion — largely as a result of the numerous difficulties that arise when interoperating between different linguistic, administrative, networking and security domains as occur over national boundaries. This section is based on the latest version of 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[2]. Text in *italics* provides discussion and clarification of the requirements.

3.1. PKI Structure

Due to certain idiosyncrasies of the grid middleware, the PKI structure within each country SHOULD NOT follow the conventional hierarchical model, but there SHOULD be a single certification authority (CA) per country, large region or international organization. A wide network of registration authorities (RA) for each CA is preferred. The RAs will 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 actual tasks of issuing CRLs; signing certificates and CRLs; and revoking certificates when necessary.

3.2. Certification Authority

3.2.1. *Computer Security Controls*

The CA computer, where the signing of the certificates will take place, SHOULD be a **dedicated machine**, running no other services than those needed for the CA operations. The CA computer MUST be located in a secure environment where access is controlled, limited to specific trained personnel and MUST be kept disconnected from any kind of network at all times. If the CA computer is equipped with at least a FIPS 140-1 level 3 Hardware Security Module or equivalent, to protect the CA's private key, the CA computer MAY be connected to a highly protected/monitored network, possibly accessible from the Internet. The secure environment MUST be documented and the documentation made available to the PMA.

3.2.2. *CA Namespace*

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

3.2.3. *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 the changes MUST be announced to the PMA for approval and approved 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.

3.2.4. *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 elements which is known **only** by specific personnel of the certification authority. Copies of the encrypted private key MUST be kept on offline mediums in secure places where access is controlled.

The pass phrase of the encrypted private key MUST also be kept on an offline medium, separate from the encrypted keys and guarded in a safe place where only the authorized personnel of the certification authority have access.

3.2.5. *CA Certificate*

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

3.2.6. *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 at least accessible via the World Wide Web, as soon as issued.

Relying parties insist that a rapid revocation mechanism is there, so we recommend that all clients update their local copies of CRLs at least once per day.

3.2.7. *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.

3.2.8. *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 secret key to sign CRLs — until all the certificates signed using the associated private key have also expired.

3.2.9. *Repository*

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

3.2.10. *Compliance Audits*

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

3.2.11. *Operational Audits*

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

3.3. Registration Authority

3.3.1. *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 CP/CPS document of the CA.

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.

3.3.2. *Name Uniqueness*

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

Some problems have been experienced with CA software when issuing new certificates with a subject name of a previously issued certificate. CAs that do issue certificates with a previously-used subject name do so only for the original subject.

3.3.3. *Records and Archival*

The RAs MUST record and archive all requests and confirmations.

3.3.4. *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 private web pages that are bi-directionally authenticated)

3.4. End-Entity Certificates

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 by 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 realise the importance of properly protecting their private data. It falls upon the user to protect his private key with a pass phrase at least 12 characters long.

Host certificates SHOULD *be linked to a single network entity. See the discussion in section 5.2 on compromise and exposure of private keys.*

4. Trust Evaluation

To establish trust, it is important to convince relying parties that each CA was founded in good faith, with clearly documented practices, and is operated in a secure environment. Each CA is required to write its own CP/CPS and demonstrate to the group that the setup is 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 recognised as a 'conforming' CA and the root certificate added to the repository.

Each relying party (RP) wants to evaluate all the CAs, either that they meet the RP's minimum standard, or that they meet an agreed common standard. The members of the CACG perform peer review on each other to establish the common standard. This allows the construction of *Trust Matrices*, the result of which is a CA *Acceptance Matrix*. All members of the CA Coordination Group are represented in the trust matrices.[2]

4.1. How to Establish Trust

To create these matrices, CA managers check each other against the agreed list of minimum requirements. Currently this requires inspection of each CA's CP/CPS by a volunteer subset of the other CAs or a trusted adjudicator. An audit of CA procedures would help, but this will be time-consuming and expensive, and none have yet been done. It is unclear if audits should be done by a third party, and who would pay for them. Evaluation of trust is a continuous and long-term process. The Global Grid Forum (GGF) [16] has established several working groups to help resolve international issues and establish policies and procedures.[3]

The evaluation process is very manual, and experience has shown that personal contacts are fundamental. All CA managers want to make the evaluation more automatic. The software that is being developed to aid this process is based on evaluation of a *CA Feature Matrix*. Each CA issues certificates according to a particular CP/CPS. Some of the features of these policies and practices have been encoded in a CA report file and the CA Feature Matrix displays the features defined in the report file.

4.2. Automatic Trust evaluation

Features of policies and practices have been encoded in a CA report file according to a basic contextual language involving key-value pairs, e.g. name = 'CERN CA'. The language is designed to enable later extension for full expression evaluation, polymorphism and matching capability of a lambda calculus to allow formal analysis, but is presently very simple. Features are evaluated relative to *rulesets*. Rules are defined within the scope of a feature's definition. Any number of rules can be defined per feature. The GGF concept of assurance levels is accommodated to allow rulesets to be defined for each level[4].

A common *default ruleset* has been defined for EDG, based on the CACG minimum requirements (see Section 3). The default ruleset would benefit from further development, but is still a useful starting point. Each VO can also define their own rules that override and extend the default ruleset, and each CA can do likewise, overriding and extending the former rulesets. This *Ruleset Inclusion Principle* extends from the general to the specific. It can be extended to users, hosts and even specific services simply by defining the appropriate ruleset. Thus a typical chain obeying this principle might be: default ruleset → VO ruleset → CA ruleset → host ruleset. It is not necessary for a typical subject to have all possible rulesets in their possession, only those rulesets in the inclusion chains that they are interested in.

For an evaluation function f and matrices A, W, R and F of acceptance levels, weights, rulesets and features:

$$A = f\left(W \times R\left(F\right)\right).$$

This assumes all possible rules are defined, whereas in practise only the default ruleset is likely to have the full complement of rules defined (which is why the evaluation remains largely an $O(n)$problem). To cope with this we add a Boolean matrix D of definition states:

$$A = f\left(W \times R\left(F\right), D\right).$$

One can delve further with the evaluations into example user, host and service certificates, and samples of issued certificates, and this is the current focus. There are other complementary approaches: for example, evaluating an XML encoding of a CP/CPS [1].

5. Related Work

This section presents work by members of the CACG performed either to improve the operation of the PKI or to explore various concerns and issues and related future developments.

5.1. Certificate Request Applets

Some of our CAs have developed Java applets to be used in the application for certificates. This is done for web-based CAs partly to overcome the need for supporting browsers via the built-in scripting APIs, but it also has other advantages.

An applet generates the keys and associated request and submits them to the CA, preferring whenever possible standard formats (i.e. PKCS[28] — in practice it is necessary and sufficient to use formats that can be read by OpenSSL). Another applet is used to download the certificate and match it with the corresponding private key (it identifies it by comparing the public keys — the 'private key' file also contains the public key). Once the certificate and key have been matched, they are exported in PKCS #12 format which can then be imported into a browser.

The applets must be signed, since they have to read andwrite files on the user's disk. Of course a signature is required anyway because the user must trust that the applets are the official ones issued by the CA. A rogue applet can get unauthorised access to the user's private key and passphrase. Alternatively, it is possible to run the applets via Sun's applet viewer and ask it to download the security policy, along with the applets and required libraries, directly from the CA. This is very easy, only one step is required to run it when the JRE is installed, but requires that a rather lengthy command line is pasted in.

Another 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 can check the format of the user's name and check whether it complies with the CA's CP/CPS. For a server/service certificate it can check whether the server has a DNS entry and whether the service is an approved service, and can inform the user directly if something is wrong. Of course it is still required that the (human) RA checks the contents of the certificate requests, just as the CA should check the contents as well before issuing the certificate.

Finally, 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.

5.2. Compromised and Exposed Private Keys

The CACG has explored the issues related to the compromise and exposure of private keys. It is widely agreed that 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:

1. If a private key can be shown to be in the possession of someone other than the user then it is considered 'compromised'. In cases where an attacker 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.

2. 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. In cases where an attacker has 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. Exposed keys discovered in the course of system administration will not normally be considered a compromise.

These are guidelines that could be used to govern the actions of the CA. In all cases, individuals might choose to take more aggressive action on their own initiative. An RA is authoritative for its users and might demand revocation of any credential it authorized. Users might choose to replace their own credentials at any time. Eventually a distillation of this definition might be introduced into the minimum requirements.

5.3. Online Certificate Services

Traditionally, grid certification authorities have been operated offline, that is, unconnected to any network. 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 is using a KCA (see section 5.3.1) and ESnet is proposing a minimum requirements profile for online services, which should allow policy management of this type of service.

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 was developed by the Center for Information Technology Integration at the University of Michigan and is distributed with several grid middleware packages, most notably the NSF Middleware Initiative (NMI)[25].

The KCA infrastructure consists of a secure server which communicates with a client application to generate PKI credentials. The client generates a private-public keypair and, using a standard Kerberos protocol service request, transmits the public-key as part of a certificate request to the KCA service. The KCA server authenticates the request through the Kerberos infrastructure and, using the client's Kerberos principal to either look-up or construct an appropriate distinguished name (DN) for the subject, creates and signs a certificate with the KCA's private-key. The resulting certificate is then returned to the client. In the default client application both certificate and private-key are stored in the client's Kerberos credential cache. The lifetime of the certificate is set to that of the enabling Kerberos token. A utility exists to extract the certificate and associated private-key from the Kerberos store to a specified location and there is also a PKCS#11 [28] library implementation which enables web browsers to utilize the certificate and key directly from the store for client authentication of secure web sessions over SSL.

As indicated above, the authentication token mapping service of the KCA is an attractive solution for sites already operating a Kerberos-based authentication infrastructure. The user is relieved from maintenance responsibility for a separate long term private key and proxy maintenance leverages the existing Kerberos infrastructure. By removing the need to run parallel registration procedures, as might be required by the addition of a traditional CA, the consequent administrative overhead, possibility of error or deliberate attack on these separate procedures is removed. Also, since the KCA issues only short-lived certificates, of lifetimes comparable to the Kerberos tickets and Grid proxies, there is no need to maintain and distribute certificate revocation lists.

When compared to a well run offline service the danger of signing key compromise is inevitably increased for an on-line service such as the KCA. Limitations in GSI complicate the effective deployment of a hierarchical CA architecture to mitigate this danger by allowing timely revocation of a compromised signing key. The minimum requirements of the CACG require hardware protection modules to be installed in such cases (see section 3.2.1), but these tend to be expensive.

Whilst, for the reasons given above, the lack of a long-term user credential is attractive, it exacerbates an issue in the context of long-running or queued jobs in the Grid: how to renew a proxy-certificate derived from a user's Kerberos token which is typically valid for only about one day. A general solution to this problem has yet to be developed.

5.3.2. *Virtual Smart Cards*

The SLAC Virtual Smart Card system [20], provides an online credential store analogous to a physical smart card. The premise is that users cannot be trusted to keep private keys secure (they may choose a poor passphrase; they may store the key on an insecure shared filesystem; they may email the key or copy it to other machines) so they should never be given access to the key. VSC can provide stronger security guarantees with a central restricted-access server than for 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 potential 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. The concept conflicts with the existing minimum requirements on the issues of the private keys being generated for the end entity by an intermediary — could the server be considered a registration authority?

— and the requirement that a user must protect his private key with a passphrase — can the VSC be considered such a form of protection?

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 production use by many Grids and application communities at the national, European and global level. The identity and authentication services provided by this infrastructure allow users and services to have just one identity credential which is accepted and trusted by a growing number of Virtual Organisations 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 long and interesting debates during the meetings of the group. The tools developed for the evaluation of trust and the various technical challenges associated with the special requirements 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 the successes of this infrastructure has already started. The expected growth of online certificate services and repositories, together with more robust and immediate 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 the European DataGrid Project. The EU CrossGrid and DataTAG projects, and parts of GridLab, joined the CACG to create a common grid authentication trust domain. The LCG high-energy physics project includes the EDG trust domain, and CAs involved in LCG were encouraged to join the CACG. 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)[14]. This new body, which is initially coordinating authentication services for the EGEE [12], DEISA [11], LCG [23] and SEEGRID [30] projects, is associated with the global Grid Policy Management Authority [19] initiative, started in 2002 to provide a common place to discuss and exchange pointers to each of the PMAs. The Grid PMA defines the International Grid Federation, which is the set of PMAs that have harmonized their policies to allow cross trust.

The policies, procedures and technical solutions developed by CACG and described in this paper, are being taken forward by the new PMA with the aim of turning this into an even more pervasive general infrastructure for authentication for e-Science.

Acknowledgements

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.

Notes

[1] Certification Authority Repository: http://marianne.in2p3.fr/datagrd/ca/
[2] Certification Authority Trust Matrices: http://www.cs.tcd.ie/coghlan/cps-matrix/cps-matrix.cgi
[3] Global Grid Forum Security Area: https://forge.gridforum.org/projects/sec

References

1. E. Ball, D.W. Chadwick, and A. Basden. *The Implementation of a System for Evaluating Trust in a PKI Environment*, volume 2 of *Evolaris*, pages 263–279. SpringerWein, 2003.
2. S. Bradner. *Key words for use in RFCs to Indicate Requirement Levels*, March 1997. RFC 2119.
3. R. Butler, D. Engert, I. Foster, C. Kesselman, S. Tuecke, J. Volmer, and V. Welch. Design and deployment of a national-scale authentication infrastructure. *IEEE Computer*, 33(12):60–66, 2000.
4. R. Butler and T.J. Genovese. *Global Grid Forum Certificate Policy Model*, June 2003.
5. L.A. Cornwall *et al.* Security in multi-domain grid environments. *Journal of Grid Computing*, 2004.
6. *CrossGrid*. http://www.crossgrid.org/.
7. DataGrid Security Coordination Group. *Security Requirements Testbed 1 Security Implementation*, May 2002. https://edms.cern.ch/document/340234.
8. DataGrid Security Coordination Group. *Security Design*, March 2003. https://edms.cern.ch/document/344562.
9. DataGrid Security Coordination Group. *Final Security Report*, January 2004. https://edms.cern.ch/document/414762.
10. *DataTAG*. http://datatag.web.cern.ch/datatag/.
11. *Distributed European Infrastructure for Supercomputing Applications*. http://www.deisa.org/.
12. *Enabling Grids for E-science in Europe*. http://public.eu-egee.org/.
13. *European DataGrid*. http://www.edg.org/.
14. *European Grid Policy Management Authority*. http://www.eugridpma.org/.
15. I. Foster, C. Kesselman, G. Tsudik, and S. Tuecke. A security architecture for computational grids. In *ACM Conference on Computers and Security*, pages 83–91. ACM Press, 1998.
16. *Global Grid Forum*. http://www.ggf.org/.
17. *Globus Simple CA*. http://www.globus.org/security/simple-ca.html.
18. *GridLab*. http://gridlab.org.
19. *GridPMA*. http://www.gridpma.org/.
20. A.B. Hanushevsky and R.D. Cowles. Virtual smart card, 2002. http://www.slac.stanford.edu/ abh/vsc/.
21. R. Housley, W. Polk, W. Ford, and D. Solo. *Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile*, April 2002. RFC 3280.
22. IETF. *PKIX Charter*. http://www.ietf.org/html.charters/pkix-charter.html.
23. *LHC Computing Grid*. http://lcg.web.cern.ch/.
24. B.C. Neuman and T. Ts'o. Kerberos: An authentication service for computer networks. *IEEE Communications*, 32(9):33–38, September 1994.
25. *NSF Middleware Initiative*. http://www.nsf-middleware.org/.
26. *OpenCA*. http://www.openca.org/.
27. *OpenSSL*. http://www.openssl.org/.
28. RSA. *Public-Key Cryptography Standards*. http://www.rsasecurity.com/rsalabs/pkcs/.
29. R. Shirey. *Internet Security Glossary*, May 2000. RFC 2828.
30. *South Eastern European Grid-enabled eInfrastructure Development*. http://www.see-grid.org/.
31. *Sun Open Network Environment*. http://wwws.sun.com/software/sunone/.
32. S. Tuecke, V. Welch, D. Engert, L. Pearlman, and M. Thompson. *Internet X.509 Public Key Infrastructure Proxy Certificate Profile*, December 2003. http://www.ietf.org/internet-drafts/draft-ietf-pkix-proxy-10.txt.

Table II. Certification Authority Root Certificate Fingerprints

CA	Country	Root Certificate Fingerprints
ArmeSFo	Armenia	MD5: 63:B3:08:9F:57:76:4A:B0:FC:D2:3D:26:15:14:CA:E7 SHA1: 9A:C4:99:EE:D5:73:3B:77:0D:04:72:69:70:66:BE:20:7F:C0:9E:01
ASCCG	Taiwan	MD5: 7E:2D:59:1F:EB:1C:93:1A:12:E5:13:65:56:16:28:41 SHA1: 46:29:F6:21:28:79:18:4D:C0:F8:BE:9D:E1:45:9C:C6:CD:D3:B2:F9
CERN	CERN	MD5: 76:B3:CF:BB:57:71:8B:80:0F:DD:D7:26:E9:18:1D:CD SHA1: 7B:86:34:13:2C:4F:39:3E:27:4C:65:84:90:13:11:4D:0A:B7:E8:96
CESNET	Czech Republic	MD5: 7F:93:95:FE:9D:76:F5:98:80:9E:5F:4B:93:1F:A8:51 SHA1: C9:00:88:24:CA:2F:99:BA:C8:CF:FB:8E:06:29:8A:38:FA:AC:AD:75
CNRS	France & Catch-all	MD5: A4:1C:C2:BB:43:34:AB:F1:C5:BD:D6:A8:50:9E:E7:2A SHA1: B9:70:D8:2C:CF:AA:DD:FC:75:07:3A:FC:8C:24:75:74:63:C3:3B:84
CyGrid	Cyprus	MD5: 72:39:00:9A:27:11:61:E2:13:01:E9:F0:3D:D0:B1:18 SHA1: 46:A3:97:D9:47:00:F4:88:C4:22:2A:30:64:53:DD:7C:2F:A4:69:F6
Datagrid-ES	Spain	MD5: DE:15:A0:5B:74:1C:53:04:7B:92:0D:79:57:6F:3F:00 SHA1: 2B:50:7A:CA:96:0C:CB:4A:8C:33:EE:2B:5E:67:0F:7F:55:C6:53:52
DOEGrids	USA	MD5: B3:76:40:75:F6:C4:BF:AF:82:CA:9A:D5:1D:FC:00:97 SHA1: 18:B9:75:4F:1D:61:AB:07:0A:8E:9B:1E:4F:A7:44:92:88:FA:D0:96
FNAL	USA	MD5: 4E:2B:6B:E1:D2:09:AE:07:B0:14:C5:5A:54:AC:10:DC SHA1: 58:16:91:F6:70:95:F8:83:C2:2C:77:BD:13:CE:47:9A:2A:F8:B6:FF
GridCanada	Canada	MD5: 3C:8F:D4:4D:BA:E2:8E:28:24:38:0F:2D:71:4C:9C:84 SHA1: 5B:A5:92:81:D1:0B:D5:75:4F:F8:D1:95:6E:B2:0E:69:32:50:C6:3B
Grid-Ireland	Ireland	MD5: 92:88:F8:93:5F:45:D8:5F:82:86:58:42:91:4F:74:A7 SHA1: 9C:C1:A7:7C:F9:C2:75:05:AD:3C:FE:3F:C8:B3:F4:76:02:25:A2:C5
GridKA	Germany	MD5: D2:FB:CC:88:63:E1:FA:83:15:64:96:96:83:22:F2:C9 SHA1: 1F:E4:41:02:EC:A7:57:D8:4A:7E:A6:EE:CC:5B:A4:19:10:57:CA:17
HellasGrid	Greece	MD5: 67:B9:5A:B5:B2:50:01:20:2C:F5:AD:D1:57:88:0D:3B SHA1: 36:12:69:64:31:35:FD:E1:FA:9B:6B:9C:4F:31:32:B5:B3:20:13:B5
INFN	Italy	MD5: 9C:3E:F4:3B:18:44:12:55:10:F3:89:C0:D5:D8:49:16 SHA1: AB:6D:CE:77:D3:5D:F0:A2:02:C8:87:4D:AE:AF:60:A8:D4:99:5C:D5
LIP	Portugal	MD5: A2:C3:F0:09:2D:61:43:BF:2F:F0:89:F7:3C:45:DE:BB SHA1: 64:1C:72:64:F8:C3:7A:BE:62:B2:C3:7E:C0:19:14:24:40:A8:78:12
NIKHEF	Netherlands	MD5: 11:7B:F0:B2:4A:2B:64:78:8A:F4:BE:87:E9:84:3D:61 SHA1: ED:C1:ED:C7:DB:46:7C:47:7F:EB:8E:CE:42:BC:12:C5:62:98:50:6E
NorduGrid	Nordic Countries	MD5: D0:BB:8D:0A:48:28:BB:8E:92:0A:D9:6C:E2:54:E5:EB SHA1: EE:49:34:9A:47:AC:0F:0D:AC:C9:31:3B:92:D0:83:A9:10:62:4C:CC
PolishGrid	Poland	MD5: AB:81:63:66:38:B2:81:B9:F8:13:11:9C:42:E6:F3:A4 SHA1: E4:18:28:73:A2:8B:03:ED:A8:A4:23:75:B2:61:D2:AB:05:1B:DB:9F
Russian DataGrid	Russia	MD5: AE:3D:F5:F2:DD:CF:B0:10:99:7A:6D:74:3C:FB:4A:22 SHA1: 7C:3E:E1:45:B8:B9:21:D6:13:1C:69:CF:B4:46:48:AD:BC:F7:7A:19
SlovakGrid	Slovakia	MD5: 05:D2:0D:A5:E0:55:CA:16:95:D4:76:C3:D3:76:8A:BD SHA1: 64:1C:72:64:F8:C3:7A:BE:62:B2:C3:7E:C0:19:14:24:40:A8:78:12
UK e-Science	UK	MD5: 32:2C:26:C7:54:47:94:51:68:4A:92:C9:0F:9F:95:E6 SHA1: 61:3F:E3:57:17:F0:4D:A0:05:CA:BB:F2:E4:BE:81:64:F1:96:02:F1