



Report from Working Group 3 of the LCG Grid Deployment Board

LCG-1 Registration, Authentication, Authorization and Security

<i>Date:</i>	4th February 2003
<i>Version:</i>	1.1
<i>Status:</i>	Draft – Interim report
<i>Author:</i>	David Kelsey



Document Log			
Issue	Date	Author	Comment
0.0	21 Oct 2002	Manuel Delfino	First Draft
0.1	25 Nov 2002	David Kelsey	Add many new sections and describe security issues
1.0	17 Jan 2003	David Kelsey	Modify following comments from GDB December meeting and start on the security model and recommendations sections
1.1	4 Feb 2003	David Kelsey	More work on security model and recommendations. Report to GDB meeting on 6 th Feb 2003

1 Introduction

The mandate of Working Group 3 (see Appendix A) is concerned with the topics of Registration, Authentication, Authorization and Security for the LCG-1 prototype. This document treats each of these in turn and makes recommendations for three different timescales; now (i.e. February 2003) as initial pre-deployment tests are started, the initial implementation of LCG-1 (July 2003) and longer term, i.e. 2004 and beyond. Recommendations are also made as to how the future work should be organised.

This document is by no means the final word on security planning for LCG. It is at best just the first step along the way. There are indeed many areas where further discussion and planning is required. One of the reasons for this is that Grid security technology and best practice will evolve significantly over the coming years. The move towards OGSA based grid services, for example, will inevitably lead to many changes. Another reason is that the various prototype grids operated to date by the current projects have been able to define rather special security environments for the relatively small numbers of users and resources involved. Site security officers have been willing to treat these testbeds as special dedicated resources with their own security constraints. This will no longer be true.

The LCG grid will be one of the first projects to deploy Grid technology on a very large production scale. This will require virtual organisations and institutes to define and agree robust security policies and procedures that will enable the building and maintenance of “trust” between the various bodies involved. In the past, distributed computing has been built mainly on one-way trust between external remote users and the large HEP computer centres. Two-way trust has, on the whole, not been required, nor have the large centres, in general, needed to agree or to trust each other’s procedures and policies. The successful deployment of the LCG grid will require many changes to the various computer centre’s policies and procedures. It is essential that the LCG community work together to build the necessary levels of trust. This will inevitably evolve slowly over many years, requiring on-going negotiation of security policy and procedures for the LCG grid, particularly following the introduction of new grid security technologies.

1.1 Definition of terms

A *Virtual Organisation (VO)* is a set of individuals and institutions that share a common goal or activity.

Authentication is the mechanism used for verifying the identity of a principal, where a principal is a named entity, e.g. a user, a system, or a service.

Authorization is the process whereby a principal is allowed to carry out or is prevented from carrying out a requested action. In the context of Grid the computing infrastructure, i.e. systems, resources and services, may also require authorization to function as part of a Virtual Organisation.

Registration is the process by which a list of authorized users and infrastructure is created and maintained.

1.2 Grid Security and the GDB working groups

The mandate of WG3 (see Appendix A) overlaps in many areas with those of the other working groups; with WG1, for example, in the choice of security middleware, with WG2 in the interaction between the management of a VO and resource scheduling and accounting and with WG4 through the security operational procedures. Wherever possible, this report takes the recommendations of the other working groups into account and aims to produce recommendations consistent with these.

WG4's mandate includes the following topic:

"Recommend a process for achieving common security operations between the participating sites, including defining and agreeing procedures for dealing with security incidents"

In a draft of their report (v 0.6), they define the need for a Site Operation Team to provide and run a Security Administration service, but state that they expect WG3 to define the detailed responsibilities of such a service. This topic is therefore also covered in this report.

1.3 Introduction to Security Middleware/Technology

The security framework of the middleware selected by WG1 is based on the Grid Security Infrastructure (GSI) implementation of the GSSAPI from Globus. A Public Key Infrastructure (PKI) is used involving asymmetric cryptography with private and public keys. The private key is only accessible to the user or machine wishing to authenticate itself while the public key is available to all. Moreover, the public key is bound to a unique identifier, e.g. the name (DN) of a user, machine or service, via an X.509 certificate digitally signed by a trusted Certificate Authority (CA). Data or messages encrypted with the public key can only be deciphered with the private key and vice-versa. Mutual authentication of the two principles involved in any grid communication is possible via protocol exchange (SSL) involving the encryption and decryption of messages using the public and private keys.

Users are able to sign-on to the grid once per session, via the creation of short-lived proxy certificates. These certificates, generated today by the *grid-proxy-init* command, and signed by the owner's long-term private key are then used to delegate the user's identity to other machines and processes allowing these to act on the user's behalf.

The current Globus toolkit implements authorisation by the mapping of a global name (DN in the X.509 certificate) to a local UNIX user account and then by the use of standard UNIX access control mechanisms. This approach, while rather cumbersome to manage, achieves the important aim of local site and system managers remaining in full control of their own resources. The local manager controls which global identity maps to which local account, via the grid mapfile, and is in control of access to the resources by these local accounts via the normal UNIX security mechanisms.

Many improvements to the Authorisation framework are under development. These will be mentioned in the later sections.

2 Definition of LCG-1 for the purposes of this report:

LCG-1 is a worldwide **data-processing infrastructure** for High Energy Physics comprising

- a **dynamically defined set of data processing resources** (computers, disks, tapes, networks, etc.) and
- the **associated** configuration and input, as well as produced, **data files** (unless exported from LCG-1) and
- the **associated** grid services, middleware, application infrastructure and services,
- which will be **provided by various institutions** in various countries
- towards the **agreed common goal** of prototyping computing for LHC experiments,
- including tests of the objective that it can be seen as a **single coherent environment from the user's point of view**
- while using **de-centralized non-homogeneous environments from the provider's point of view**,
- **built and operated** according to an architecture and standards agreed upon by the LCG project

LCG-1 will be:

- **utilized** by LHC experiments, and possibly other projects for test purposes, to perform a portion of their data processing work load and to test the viability and scalability of the prototype for LHC computing
- **built up** during the first half of 2003 and has a major milestone of **start of operations in July 2003**.

3 The LCG-1 Security Model

This section contains an overview of the proposed security model for LCG-1 in terms of the topics of the mandate of WG3. This is based on user and site requirements, as collected and documented; for example, in the EDG Security Requirements document (D7.5) and the RTAG4 HEPCAL common use cases. This section contains requirements and assumptions for the implementation of LCG-1 security. The recommendations in this document are made to achieve the implementation of this model.

The scale of the number of users and sites involved in LCG dictates that the overall long term model must be that sites negotiate the provision of and control access to resources and services in terms of VO's rather than individual users. VO's in turn register their users, authorise their access to the VO resources and revoke this authority when a user leaves or following certain security incidents.

To achieve the required level of trust between sites and VO's, it will be necessary for the VO to be "real" in many senses rather than virtual. To achieve this it is recommended that VO's are managed and operated by the main home site of the experiment. For LCG this means that CERN must take responsibility for the appropriate management of the 4 experiment VO's.

Access to site resources may be based on the user's membership of more than one VO. Priority may be need to be given, for example, to local users. This can be achieved by the local users being members not only of the main experiment VO, but also of a local site VO.

None of this removes the requirement that sites still definitely need full information about individual users, both in the form of audit trails and personal information in a locally held database (or local



copy), and the ability to control local access to resources not only in terms of VO's but for specific users.

3.1 Registration

Before joining LCG-1, the user must generate a private/public key pair, and obtain a signed X.509 certificate from their local approved national Certificate Authority (CA). Many of these CA's are run for a much larger community than just HEP, so possession of one of these certificates is not sufficient to prove authority to use LCG-1 resources. In future, additional types of authentication will also be supported, for example the authentication via Kerberos with the automatic creation of short-lived proxy certificates for grid use.

LCG-1 will need to provide and run a catch-all CA, willing to sign certificates for users who have no access to a national CA. Today in DataGrid this function is provided by the CNRS CA, but there is no agreement to continue to do this for LCG.

Users then register once, and **only once**, to be granted authority to use LCG-1 resources.

They complete a single form (preferably electronic) providing all the required personal information. By this process the user requests to join the LCG-1 VO, one or more of the experiment or special VO's and to join the VO of his/her home institute, if it exists. The initial registration may also include the request to join specific groups and/or sub-groups within a VO and/or to request the authority to act with a specific role or capability.

In the long term it is desirable that the registration procedures are managed and operated as part of the normal user registration process by the CERN experiment secretariats.

Users will be required to sign an Acceptable Use Policy (AUP) or Usage Guidelines document. The aim is that this is a **single** document, which may need to refer to the official rules and policy documents of each of the LCG-1 sites. EDG has such a document (<http://marianne.in2p3.fr/datagrid/documentation/EDG-Usage-Rules.html>) that also includes implicit reference to the rules of all sites. We should start from here and see what a) sites will agree to and b) what users need to know before using LCG-1.

An important feature of the registration approval process is that robust checks will need to be made to confirm that the user is who they say they are and is entitled to have access to the resources, VO's, groups and roles requested. This will require approval by each of the VO's requested via a distributed VO-specific Registration Authority (RA) before the user registration is completed.

In the longer-term it is desirable that machines and services are also registered for authorisation to run within a VO, but this is some way off.

Once a user has successfully registered, the user accounts, either fixed or dynamic pooled accounts as the site requires, will be created at each of the sites used by the VO of which the user is now a member. This will happen automatically, within an agreed timescale, without any interaction between the user and the sites. Any queries or problems will be fixed by communication between the site and the administration of the VO.

3.2 Authentication

A list of approved and trusted national CA's will be maintained by LCG-1. It is expected and very much hoped that there will be a (small) number of global Policy and Management Authority bodies (PMA) run on a regional/continental basis for all HEP-related CA's used by LCG thus removing the need for LCG-1 to run its own PMA.

Certificate Revocation Lists (CRL's) must be maintained by the CA's and distributed to each of the sites regularly (at least daily). There is perhaps no need to revoke more quickly than this, as urgent removal of rogue users will be achieved by revocation of authorisation.

Users require the feature of single sign-on to the Grid with the generated proxy credentials being valid for a reasonable time (the default today is 12 hours).

In the longer term, LCG-1 will have to consider the possible use of additional, alternative authentication technologies, which may be required by at least some of the sites. The ideal aim is, of course, just one method of authentication for LCG-1, but if more than one is required the list should be kept to an absolute minimum. The requirement of user single sign-on will at this point, of course, no longer be possible.

Users need to be able to run long jobs, i.e. those that will run longer than the lifetime of the grid proxy certificates (default life 12 hours). This requires some form of automatic credential renewal by the job or job manager on behalf of the user. DataGrid WP1 WMS has a solution today based on the MyProxy credential repository but this is far from ideal. Development will be required to improve this situation.

3.3 Authorisation

Users require authorisation to perform grid tasks. This includes control of access to many different types of grid objects; compute elements, storage elements, files, queues, information services, etc. This may be coarse-grained, in the sense of control of access to the overall service (e.g. an SE), or fine-grained in the sense of access control based on a particular file residing on the SE.

VO's require the ability to dynamically control the roles that a user may be granted and/or the groups or sub-groups to which a user may belong.

Sites must always have final control over the access to their own resources. Individual specific users can be restricted even if the VO authorises them.

WG1 has defined in its report that the initial implementation of authorisation will be based on the EDG VO/LDAP directory and the associated tools to aid automatic generation of the grid mapfile (mkgridmap).

Future developments in this area include the EDG security components (VOMS, LCAS, edg-java-security, etc) and the Globus Community Authorisation Service (CAS). These will be needed to allow more dynamic, role and group-based, and fine-grained access control to grid services and objects.

4 The LCG-1 Security Policy

A full security policy needs to be developed, discussed and agreed. This document, which should be completed during 2003, should include the aims of LCG-1 in terms of implementing and operating a secure grid and the methods by which this will be achieved. The user, VO and site responsibilities should be described and implications and actions that will be taken if users, VO's or sites do not abide by the agreed policies and rules. This document should be developed in parallel with the Usage Guidelines document

5 WG3 recommendations

These are described for each of the main topics in 3 different timescales; now (i.e. February 2003) as initial pre-deployment tests are started, the initial implementation of LCG-1 (July 2003) and longer term, i.e. 2004 and beyond. These are very much constrained by what is available today and how much work will be required to develop new procedures. Recommendations are also made as to how the future work should be organised.

5.1 Registration Recommendations

	Now	July 2003	2004
From where do users obtain a certificate?	National EDG/DOE approved CA	National EDG/DOE approved CA	National EU/North America/Asia PMA approved CA
And host/service certs?	National EDG/DOE approved CA	National EDG/DOE approved CA	National EU/US/Asia PMA approved CA
What if no national CA?	Ask CNRS (but subject to RA agreement)	CNRS or LCG-1 catch-all CA (CERN?)	LCG-1 catch-all CA required (CERN?)
Where does user register with LCG-1?	Sign the EDG usage guidelines (web form)	Sign the LCG-1 usage guidelines (web form)	Sign the LCG-1 usage guidelines (web form)
Which version of AUP or Usage Guidelines?	The current EDG guidelines	LCG-1, but based closely on the EDG guidelines	New LCG-1 guidelines brought into service.
Who can join LCG-1?	Anyone with a certificate from an approved CA	Anyone with a certificate from an approved CA	Anyone with a certificate from an approved CA
Is there an LCG-1 VO? This is the collection of users who have signed the LCG-1 Usage Guidelines.	No. Use the EDG Guidelines VO. Users of pre production LCG-1 have to sign the EDG guidelines	Yes – to be setup and run by LCG at CERN	Yes. At some point, this should be maintained by data extracted from the central CERN databases.
Where is the user's registered personal data stored?	In the EDG VO database	In the LCG-1 VO database	At some point, move to the data being extracted from the CERN HR/Experiment/Computer databases
Who can join an experiment VO?	Anyone confirmed to be a member of the collaboration. Checked by EDG VO manager	Anyone confirmed to be a member of the collaboration. Checked by EDG VO manager	Anyone confirmed to be a member of the collaboration. Checked by distributed experiment RA's
Definition of personal data required in the VO database (superset of site requirements)	Use the currently collected EDG information	Minor changes to the EDG schema but now running on the LCG-1 VO	New LCG schema brought into operation
How are user accounts at each site?	As today in EDG. Manually via direct individual contact.	Aim for at least some level of automation. Leased pool accounts will help if site willing to do this.	Fully automated procedures. No user intervention required.

5.2 Authentication Recommendations

	Now	July 2003	2004
Which CA's does LCG trust?	The list of EDG/DOE approved CA's. (Note that Sites are always able to define their own list – but this is not encouraged)	The EDG/DOE list	The list approved by the PMA's (EU/North America/Asia pacific)
Support of online Kerberos CA's?	No	Yes (EDG will work with FNAL and CERN on this)	Yes
Supported Authentication mechanisms?	Standard GSI – with proxy certs generated from user-held long-term key pairs.	Standard GSI and explore methods to better secure user-held private keys (VSC etc.)	Standard GSI, more secure private keys (when do real smart cards become viable?) and perhaps site-specific authentication call-out, e.g. Kerberos? (but try to avoid if possible?)
Online credential repositories	MyProxy service for credential renewal (WMS)	MyProxy + VSC	Work required to improve the security of online credential repositories (in collaboration with EDG, DOE, GGF, ...)
Certificate Revocation	CRL's produced by CA's. Sites must fetch CRL at least once per day.	CRL's produced by CA's. Sites must fetch CRL at least once per day	CRL's produced by CA's. Sites must fetch CRL at least once per day

5.3 Authorisation Recommendations

	Now	July 2003	2004
Who runs the Experiment VOs?	EDG – run by Nikhef and the existing experiment VO managers.	EDG – run by Nikhef and the existing VO managers	LCG should take over the running of the VO databases after EDG has finished and run it on machine(s) at CERN with new VO managers
Specialised VOs? e.g. other communities, development and testing	Negotiate with EDG to see if can be added	Run by LCG	Run by LCG
Authorisation middleware? (see WG1)	EDG VO/LDAP and mkgridmap tools	LCG VO/LDAP and mkgridmap	VOMS/LCAS/LCMAPS GACL/SlashGrid (and investigate CAS?)
Access to user personal data for audit and incident tracking	EDG VO is currently world readable	LCG VO/LDAP database (need to investigate security and legal aspects of data access)	LCG VOMS database (need local mirrors) (But who can read?)
Do sites create a local site VO?	No	Perhaps?	Yes? Needs more investigation into the use of this (method of proving affiliation to home site)
Do LCG-1 VO's need to authorise infrastructure? (i.e. machines, services)	No	No	Perhaps? Needs more investigation

5.4 Security Recommendations

WG3 needs to survey Site security requirements to see where agreement can be reached and where there are still differences of opinion. The GGF Site-AAA requirements research group and the PPDG-SiteAA project are both active in this area.

http://www.gridforum.org/2_SEC/SAAA.htm

<http://www.ppdg.net/pa/ppdg-pa/siteaa/>

Dane Skow, the Computer Security Executive (Deputy) at FNAL, is playing a leading role in both the PPDG and GGF activities. To speed up the gathering of information Dane was asked if he was willing to extend these surveys to the LCG-1 sites not already included. He has agreed to do that and a questionnaire (see Appendix B) was sent out during January 2003 to the Security contacts for each of the LCG-1 Tier-1 sites. Feedback from this process can then be used to improve the questionnaire before contacting the remaining LCG-1 sites. This survey will report back on site issues and requirements, including incident handling and firewall issues.

Recommendations for now (Feb 2003): Continue to do what we do today in EDG and DOE testbeds. Firewalls will be configured as required and incident handling will use the existing network security reporting channels. It is however likely that the site survey will result in the recommendation to create channels of communication between all of the LCG-1 site security contacts.

The answers from the survey should then be used to determine future developments for July 2003 and 2004 and beyond.

5.5 Future organisation and future work

As stated in the Introduction, this paper is only the first step along the way of defining the security plan for LCG-1. The work very much needs to continue, and there are still many issues to solve.

WG3 recommends that a "LCG GDB Security" working group should be formed to continue the planning, implementation and deployment during the next 12 months. The membership of this group should be drawn from the: Experiments (VO's), Site managers, site security contacts, security middleware experts, CA managers, and LCG management. The ongoing need for such a group should be reviewed early in 2004.

6 Appendices

6.1 Appendix A: Reproduction of the instructions given to the Working Group by the GDB (extracted from <http://lcg.web.cern.ch/LCG/PEB/gdb/WG3/>)

WG3: LCG-1 Registration, Authentication, Authorization and Security

Mandate:

Recommend how common Registration, Authentication, Authorization, and Security policies and procedures should be achieved to ensure integrated availability of LCG-assigned infrastructure and resources to valid users. Identify technical issues for Security, Authentication and Authorisation that need to be undertaken by the LCG to meet the deliverables for LCG-1 in July 2003, e.g.

- developing the model for simple user registration, and maintenance of the user database(s)
- deployment of a common model for mapping VO users to local users,
- implementation and operation of the accountability functionality that satisfies the security needs of the sites.

Authentication:

- Survey the existing agreements and trusts for CA/RA between EDG, EDT, DOE, etc. for incompatibilities and determine a procedure for deciding if these are sufficient for the needs of the sites participating in LCG-1. If so, recommend a process for agreement by the sites. If not, define the procedure to be followed to add the missing sites.
- Identify the constraints from participating sites on authentication and accountability.

Authorization:

- Recommend steps towards acceptable use policies for LCG-1. Ideally the goal should be that a user need only sign once. Define a process for reaching agreement by all participating sites that covers all their requirements for acceptable use of resources.
- Recommend a process and tasks to define and deploy the model for authorization that meets the constraints of the participating sites, and that satisfies the requirements of the experiments. This should include appropriate layers of delegation of authorization.

Security:

Survey the security requirements of the participating sites and recommend activities and policies to be developed and implemented by the LCG.

Members: Manuel Delfino (chair), David Kelsey (tech. assistant), Thomas Kachelhoffer, Randy Sobie, Vicky White



6.2 Appendix B – Questionnaire sent to Site Security Contacts

This was distributed to LCG-1 Tier-1 security contacts by Dane Skow (FNAL) on 24th January 2003.

- 1) Do you require agreement to an Acceptable Use Policy for users of your computers? (Could you give a link to it so we can compare them?)
 - 1a) Will you accept the principle of a single "LCG Usage Guidelines" document, which then refers, either directly or indirectly to your own Acceptable Use Policy? If not, please state why (e.g. special legal constraints)?
- 2) Is this agreement required to be in person? via written signature? Will you accept 3rd-party assurances that this has been collected (e.g. If the experiment secretariats were to do this for you) or do you require a copy?
 - 2a) If you will not accept these 3rd-party assurances, please state why and do you see any other way of achieving automated user registration?
- 3) Do you require individual accounts per person? (are shared password accounts allowed?) What registration information is required?
 - 3a) Are you willing to use pooled accounts (per VO) or do you require a real user "name"?
- 4) What requirements do you have on authentication methods for login access? for file transfer? other categories?
- 5) Do you block network access to certain TCP or UDP ports? Is this list public? Is this list available to this group?
- 6) Are you familiar with the Globus Grid Security Infrastructure?
- 7) Will you make block grants of CPU-time for others to delegate use? Do you have restrictions on who could use it? Do you have restrictions on what applications they run?
- 8) What requirements do you (expect to) have for incident response follow-up for Grid access to your resource (in the event of a case of misuse? in the case of an external attack?) What expectations do you have of Grid sites in investigating incidents?
- 9) Do you (expect to) have special requirements on machines that offer Grid services (e.g. special network restrictions, executable restrictions (sandboxes), ...)?
- 10) Do you have expectations on logging requirements of (Grid) accesses of your peer Grid sites? What access do you expect (e.g. for incident handling) to those logs? to the keepers of those logs?
- 11) Do you have other concerns about Grid computing that you would like to raise (or discuss) with this group?
- 12) Would you like to participate in discussions of Grid Security with your LCG peers? Do you feel you have a channel to do so? Should this be ongoing ?