ESP-GRID - Evaluation of Shibboleth and PKI for Grids

Draft

Requirements gathering exercise: (1) Use cases for a generic grid

Document History

Version	Date	Comments
0.65	27 April 05	Lots of rel. minor changes following emails from Von Welch and Bruce Beckles. Major change is to the definitions section, following criticisms from the aforementioned, plus others.
0.6	25 April 05	Moved the example stories to an appendix as they are clearly (from feedback) far too distracting! Modified some of the stories after Mike Fraser's feedback. Also changed ASP->SP after discussion with Christian.
0.5	22 April 05	Moved the example stories forward in the document (was sect 3, now s2). Also, incorporated some of Shawn Mullen's ideas on HIPPA/confidentiality. Also moved final column of main table to a new section.
0.4	8 April 05	Set of v. useful comments from Ivo and Francisco. Plus, I added the scenario of the 'forever' grid job/routine and wrote the silly examples in section 3.
0.3	7 April 05	Not quite finished – out for comments though. Maybe a problem with layout of the main table.
0.2	1 April 05	Little bit more work. Possibly did not circulate
0.01	30 March 05	First uncompleted draft. Circulated to Alun and Ivo.

Contents

1. Introduction

1.1. An appeal

The authors of this document appeal strongly for feedback and for criticisms of our use-cases. We would also like to hear examples of other use-cases that we may not have considered.

1.2. What is this document?

This document is the first step in establishing the types of actors and some example use cases, and then requirements for access management and security in a 'generic grid'. We use the term *generic grid* to denote something which is a grid (defined below) but which does not (as yet) mandate any particular technology or middleware. Once the general use cases are established that reflect the types of users that a generic grid would support, and possibly some of the scope of activities within a grid, then the requirements for the middleware/access management and security can be generated.

1.3. What is a grid?

1.3.1. Other people's definitions

> *An environment in which individual users can access computers, databases and experimental facilities simply and transparently, without having to consider where those facilities are located. [RealityGrid, Engineering & Physical Sciences Research Council, UK 2001]* http://www.realitygrid.org/information.html

> *A means of network computing that harnesses the unused processing cycles of numerous computers, to solve intensive problems that are often too large for a single computer to handle, such as in life sciences or climate modeling. http://www.consultingtimes.com/glossary.html*

After admitting that there is a short answer and a very long answer, the GridCafé web pages at CERN (*http://gridcafe.web.cern.ch/gridcafe/whatisgrid/whatis.html*) say that:

> *The short answer is that, whereas the Web is a service for sharing information over the Internet, the Grid is a service for sharing computer power and data storage capacity over the Internet. The Grid goes well beyond simple communication between computers, and aims ultimately to turn the global network of computers into one vast computational resource.*

Wikipedia (http://en.wikipedia.org/wiki/Grid_computing on 29 March 2005), described grid computing, thus:

> *Grid computing offers a model for solving massive computational problems by making use of the unused resources (CPU cycles and / or disk storage) of large numbers of disparate, often desktop, computers treated as a virtual cluster embedded in a distributed telecommunications infrastructure.*

The same article later asserted:

> *Grid computing involves sharing heterogenous resources (based on different platforms, hardware/software architectures, and computer languages), located in different places belonging to different administrative domains over a network using open standards. In short, it involves virtualizing computing resources.*

Ian Foster (with Carl Kesselman) updated his previous definitions of a grid in 2004. It should be noted that Foster has also come up with checklists and other, more lengthy text to explain what is a grid. Foster and Kesselman stated:

> *We define a Grid as a system that coordinates distributed resources using standard, open, general-purpose protocols and interfaces to deliver nontrivial qualities of service.*

1.3.2. Our definitions

For the purposes of this document, we take much of the spirit encompassed in Foster and Kesselman's definition, but find the phrases "standard, open" and "nontrivial qualities of service" laudable but not necessarily defining terms for a grid. We therefore define a grid as:

> ***A set of networked computers and/or other devices, including remote instrumentation, that have been made available so that their operation can be shared. The sharing of these resources must be via an agreed set of protocols.***

Foster and Kesselman's "system" is an object because it is identifiable by the agreed set of protocols. Any grid system which the ESP-GRID project produces will use "standard, open, general-purpose protocols", but it is possible that other grids may use proprietary code and standards, as long as all components of the grid use the same protocols. However, for resources that are geographically remote and non-contiguous in network terms, the feature of the set of resources that conveys the essence of being a grid is the common protocols (or possibly middleware).

N.B. For the purposes of the ESP-GRID project, **we must also assume that the 'generic grid' is of a mixed economy** – i.e. that commercial, academic and non-profit use may co-exist within the same grid. This means that we must consider grids where detailed accounting must be possible. However, this does not need to affect the definition of "a grid".

1.4. What this document is not

This document is not about building requirements for access management and security. The approach taken with this document is to try to capture some scope of a generic grid, identify the basic actors in such grids using example use-cases and from there to open the way for thinking (in documents) regarding general access management and security requirements.

2. Use-cases for grids

2.1. Introduction

Please note that section 2.3 addresses the use-case scenarios from the point of view of a user, rather than the technology or machines involved. The use-cases in section 2.3 do not consider issues such as personal privacy (as in service providers knowledge of who is the end-user) and data confidentiality (as in service providers being able to steal sensitive or confidential data). These issues are considered in section 2.4 on page 9.

See section 2.2 (on page 4) for short definitions of the terms used in the next section.

The accompanying/later document *Requirements gathering exercise:(2) Authentication, authorisation, accounting and security* goes on to consider some of the issues of authentication, authorisation, accounting and security (AAAS) that arise from such use-cases and the types of grids that are proposed in this document. It should be noted that the subsequent document is fully dependent upon the contents of this present document and any changes to the use-cases may affect the AAAS findings as laid out in that later document.

As a point of interest with which to explore the actors and to test the generalised use-cases given in section 2.3, Appendix 1 on page 10 gives some example user stories.

2.2. Terms defined

AuthN Authentication.

AuthZ Authorisation.

Grid AM service Grid access management service (left undefined further).

Grid resource node Any computer or instrument connected as part of the grid that is available for grid use.

GRID-SYS Grid Infrastructure System Administrator (or similar role).

Identity provider Authentication (and possibly authorisation) service run by an organisation to which the user belongs. The grid, or the *grid AM service* may trust this identity provider to perform the authentication task and it may also trust this provider to supply reliable authorisation/status information.

Primary grid service A grid infrastructure service (e.g. grid cluster, resource broker, access management point etc., as opposed to other ‘services’ which may run as applications and which may use the primary grid services.)

PUA Power user that does not care which *grid resource node* is used to run his/her job.

PUDS Power user developing an application service.

PUS Power user requiring specific *grid resource nodes* upon which to run jobs.

Resource broker Some kind of service running on ‘the grid’ that accepts jobs from users and *SP*s and allocates those jobs to individual *grid resource* *nodes*. It may also play a role in accounting.

Note: This document seeks to avoid mandating architectures or middleware technology. Therefore, the resource broker could also be taken to mean the first node to which the user submits a job. That primary node could also negotiate with other nodes in order to run jobs.

Secondary resource
Computer that is not dedicated to only grid use. Its primary purpose may cause it to be under-utilised at certain times and (as a secondary purpose) it can be used as a grid resource.

SP
Service Provider.

SEU
Service end-user.

TPB
Third party beneficiary of grid processing. This entity (individual or organisation) is not a user (PUA, PUS, PUDS or SEU) but a grid user may invoke a grid procedure that processes that entity’s data.

2.3. Types of grid users and very basic example scenarios

We propose the following types of grid users and give some example use scenarios. Terms in italics are defined in the previous section.

Type of user	**Typical characteristic**	**Example use-case / scenario**
Service end-user (SEU)	No computing expertise	PhD Biologist submitting large data sets for processing or Humanities researcher asking very complex questions of a service (e.g. requiring complex textual analysis). or User or organisation receiving regular output (without necessarily sending input) e.g. the BBC or Meteorological Office receiving bulletins from a ‘Weather’ SP.
	AM/security characteristic:	SEU does not need to be ‘known’ by the *grid AM service* (as the grid trusts and accounts the *SP* not the user). SP may need to authN/authZ and account for the user.
Power user agnostic of *grid resource node* (PUA)	Develops programs and data but does not care where processing takes place	Technical expert programmer supporting end-user. Submits the programs and data to a *resource broker* or primary node, which, in turn, submits jobs to (other) *grid resource nodes*. The PUA does not care which resource takes on the job. Example: Takes data from PhD Biologists as there is no service available for their needs. Packages data and algorithms and submits these to the grid for processing.
	AM/security characteristic:	PUA need not be ‘known’ by the *grid AM service* (but some sort of mapping to a billing account may be necessary). It is likely that the grid AM service may need status information from an *identity provider* (for authZ purposes).
Power user requiring specific *grid resource nodes* (PUS)	As PUA but may have more platform etc. dependent expertise and some sysadmin expertise	As above (PUA) but PUS does not wish, or cannot, use a *resource broker* (in its normal method of operation). The PUS writes specifically for jobs to be run on defined grid nodes. This could involve interaction with a resource broker, but for accounting purposes only. Example 1: (The example of an expert serving the needs of PhD Biologists or Humanities researchers fits equally well here). Example 2: PUS has a never-ending project that calls a grid-connected telescope studying sunspot activity. PUS has to be specific about the telescope and s/he is also driving a project that needs to keep running and not be seen as a discrete (set of) job(s) that has one output.
	AM/security characteristic:	PUS may or may not need to be ‘known’ by the *grid AM service* (but some sort of mapping to a billing account may be necessary). All of the text regarding AM/security for PUA is equally valid here. However, in addition, grid node owners may wish to have a direct authN/authZ (and accounting) relationship with the PUS.

Type of user	**Typical characteristic**	**Example use-case / scenario**
Power user developing a service (PUDS)	As PUA/PUS but developing expertise like SP	As PUA or (more likely) PUS where the user wishes to allow the developed application to run as a service for SEUs but the service is still in development.
	AM/security characteristic:	As for PUS or PUA, but moving into arrangements like SP (see below). May need to begin interacting with and accounting for SEUs in an experimental manner.
Service Provider (SP)	As PUA/PUS but has expertise in authZ and possibly identity management	SP provides a user interface (possibly via web, not necessarily via grid middleware) for SEUs. The SP interfaces directly with SEUs and then adopts a role as PUA or PUS in order to execute the processing job. Example 1: Accepts large spreadsheets or XML files of data from PhD Biologists (or digital texts and complex textual analysis questions from humanities researchers). Example 2: A 'Weather service' SP runs constant 'chains' of jobs on the grid that call upon satellites and weather stations for 'moment in time' data. Grid jobs compute predictions and reports to present or send to SEUs. The SP may need to identify or authZ the SEU for access or accounting purposes. The SP then submits the 'job' to the grid, possibly via a *resource broker* or possibly directly to particular grid nodes. The SP collates the returning output and sends or presents it to the user.
	AM/security characteristic:	SP may be trusted to provide services only to those supposedly authorised to use the grid. SP may need to identify (authN) SEUs but will probably need to recognise status (authZ). SP will need strong authN between it and the *primary grid service* or *grid resource nodes*. Accounting may be required between the grid resource nodes (or primary grid service) and the SP and between the SP and the SEU (although this latter requirement may not need to be met using grid middleware).
Infrastructure sysadmin (GRID-SYS)	System administration of grid nodes with possibly infrastructure delivery and security expertise	A GRID-SYS may manage dedicated *grid resource nodes* (including clusters) and any grid system objects such as resource brokers, authN, authZ or accounting points. As well as possibly managing a resource, a GRID-SYS is likely to be responsible for (and expert in) security and access management. A GRID-SYS may be the resource manager of a node that accepts jobs (from PUAs) from the resource broker, or of a node that may authenticate or authorise PUS users directly where they wish to be specific and use the GRID-SYS's resource without any involvement of the resource broker. A special type of GRID-SYS is someone who hosts a grid resource node for a particular SP, or a set of SPs.
	AM/security characteristic:	A GRID-SYS may need to authenticate directly to particular *grid resource nodes*. However, in theory, it is possible that s/he may authenticate elsewhere and the node computer may trust that external authentication point (or identity provider). [This may be difficult to accept in these days where direct (root) access for sysadmins is the norm, but it would seem that there is no compelling reason for this to remain the primary system of access in the future].

Type of user	**Typical characteristic**	**Example use-case / scenario**
Third party beneficiary TPB (non-user)	Person or organisation who does not interact directly with the grid but where his/her/its personal data are being handled on the grid	Data belonging to or pertaining to a TPB may be handled by one or many grid nodes. These data may be required to be guaranteed to be confidential and the TPB may require anonymity. Example: A SEU could ask an SP for a TPB's records to be processed and for the SEU to receive the results of the processing. Irrelevant data concerning the TPB may be required to be kept from the SEU and all other grid users, owners and administrators. The TPB should be anonymous or untraceable by other grid users, owners and administrators.
	AM/security characteristic:	The TPB typically does not interact with the grid (whereby s/he would become a SEU) and therefore no direct AM may be required, except for interaction with the database that holds the confidential data (and this may not be considered a grid interaction – merely database authN/authZ). However, there is a possible security characteristic in that the TBP's data and identity may have to be kept secret from other grid users.

2.4. Models of grids and grid resources

The following is a (non-exhaustive) list of types of grid resource and models of grid upon which grid computing may be possible. N.B. All may be possible on the same grid, and examples from section 2.3 may be applicable to all.

A) Dedicated primary grid service
(e.g. compute cluster, data cluster)

B) Voluntary secondary resource, actively monitored by resource owner.
Resource owner deliberately makes resource un/available and may choose whether or not to run grid jobs on an individual basis.

C) Voluntary secondary resource operated blindly by resource owner, possibly with dedicated, secure, ring-fenced sandpit within the system that defers to end-user activity.

D) A no-trust, no-accounting grid (subset of *C)*, above). Each node has a secure sandpit and the owner allows anything to go on there. All users are authorised to use it.

2.4.1. Notes/examples:

SETI@home and climateprediction.net should be examples of *B)* above as they could theoretically be managed by the resource owner and be actively selected. However, as most workstation users completely trust the programs, they may be behaving more like *C)*, except that the processing is not ring-fenced and secure.

No further detail is attempted here as this document attempts to be neutral in terms of architectures and technology.

2.5. Privacy and confidentiality

Running alongside each of the use-cases above are another two dimensions. The first is the need for privacy/anonymity and the second is the confidentiality of the data and/or algorithms.

2.5.1. Privacy

In any of the use-cases listed in section 2.3, the identity of the end-user may need to be protected. Grid nodes and services may care *what* the end-user is, but may not care *who* is the end-user. Clearly this is easier to achieve if a trusted third party (e.g. a SP) is submitting the grid job(s).

2.5.2. Confidentiality of data and/or algorithms

Again in any of the use-cases listed in section 2.3, the data and/or algorithms being processed may either be sensitive (e.g. medical records) or confidential (e.g. of commercial importance). Users may need either contractual guarantees that data or algorithms cannot be stolen or observed by an ‘unauthorised’ entity, or for this to be technically unfeasible.

1. Appendix one - Some example use-cases (end to end stories)

This section contains some story-line cases with which to illustrate the generalised use-cases contained in section 2.3 on page 6. **This is a (near) trivial example section, and is merely for feeding the discussion regarding the broad use-case definitions within section 2**. This section of the document does not attempt to encompass the broader issues and the many types of users. However, section 2 attempts to do this. Abbreviations used in this section are introduced within section 2.

a) A humanities researcher (SEU) submits a text document containing metadata and a set of video data to a grid SP and asks for a very complex multi-factor analysis involving the text and the video data.
The SP needs to know that the user has the correct privileges to use the service and must find out that he or she is a member of the UK academic community and already holds a degree.
The SP also needs to know to which organisation (department and institution) the user belongs in order to bill (*charge financially*) that organisation.
The processing requires the use of three grid nodes. The SP submits the job and auditing/tracking metadata so that the grid nodes may bill the SP.
Periodically the grid nodes bill the SP and the SP has its own charging mechanism for billing the humanities researcher.

b) The BBC Weather Unit in London (SEU) registers to receive hourly weather data output from the profit-making UK Meteorological Grid Service. The UKMGS has a data cluster and compute cluster of its own, but regularly has to purchase processing power from nodes on the UK e-Science Grid. It also demands specific output from several grid-enabled satellites and government-run weather stations. This is done in an automated but unpredictable way (e.g. for a particular combination of temperatures and pressures, the UKMGS jobs may ask for radar data for – unpredictable – regions around the British Isles).
Each grid node 'called upon' by the UKMGS jobs charges the UKMGS for the processor or instrument time used.
The UKMGS puts its output data on secure web sites for the BBC Weather Unit to pick up. The BBC pays a standard fee for the service, but occasionally will pay more for specific requests for 'unusual' data, such as "What will the weather be like for the England game on Tuesday in the World Cup finals in Munich?"

> b)i The UKMGS is very protective of the algorithms that it has produced for predicting the weather. It needs to be able to run its jobs and to have a guarantee that the owner (or other users of) the external grid node will not steal the data or the algorithms. Ideally, the UKMGS would like this to be technically unfeasible.

c) The Smalltown Medical Center receives an unexpected patient when the President of the USA visits town. The President supplies a sample of blood for which the Smalltown doctors need to scan for a variety of pathogens, toxins and other markers that could be indicative of his symptoms. The analysis of the blood and the cross-checking with other data held in secure databases concerning the patient is very processor-intensive. Therefore the Medical Center (SEU) uses a grid service provider (SP) to process the job. This SP must be able to run jobs which can query the secure database so that only *positive results* are reported to the Smalltown doctors (i.e. the Smalltown doctors must not be able to know which diseases the patient has suffered with in the past, unless they are clearly relevant to the analysis as has been indicated by the algorithm of the grid job). The SP may, or may not, own the secure database. Whilst the data are being moved around the grid, the privacy/anonymity of the patient must be guaranteed as well as the confidentiality of all data that have proved irrelevant to the final findings.[1]

d) A biologist researcher needs some very processor-demanding work to be performed for a statistical analysis of a very large data set that s/he has collected. His/her IT support specialist is able to write a program to perform the work but must submit this program and data to the grid for the job to be performed.
The IT specialist already has the access credentials to be able to run jobs on the grid, but has to guarantee to a *grid AM service* or auditor that the researcher is also privileged to benefit from such work.
The IT specialist submits the job to the grid and does not care upon which grid node the processing takes place (acting as a PUA). The job is completed and the specialist picks up the results and passes them on to the researcher.
The grid node (or resource broker) may demand payment for the use of the resource, but the biologist is part of a community that should receive such services without charge. This is expressed to the grid node or resource broker.

d)i The biologist thinks that s/he may be physically attacked by people who morally oppose the nature of his/her work and wishes to remain anonymous and untraceable.

d)ii The biologist believes that s/he is close to finding a cure for HIV and does not wish for the Nobel Prize to go to anyone else, should they see some of his/her data or see the way s/he is interrogating it. Therefore, s/he needs the data and algorithms to remain a secret from other grid users.

e) A theologian has a very complex textual analysis of a great number of published versions of the bible. The question is too complex for any of the available text mining services that are currently resident on the grid and so the researcher has to have a developer design a program to carry out the analysis.
The developer (acting as a PUS) knows of a data cluster that already contains copies of these versions of the Bible and so writes a program or ‘job’ which needs to specifically access this data cluster.
At some point the developer has to prove that the theologian is privileged to make use of these grid resources.
The job is run and the theologian’s university department is billed for the use of the grid resources.

e)i The theologian is aware that his/her research is highly controversial and therefore wishes for his/her identity to remain secret. S/he may still need a mechanism of ensuring that the SP is paid for its services

e)ii The developer is thinking of making money from his/her algorithms and wishes for them to remain secret.

f) Later, the PUS developer (from example e)) realises that there are many researchers that need similar jobs carrying out. Therefore s/he develops a web interface, that includes a billing mechanism, for researchers (SEUs) to use to choose texts and cross-referencing queries.
The developer becomes a private company and pays for grid membership so that s/he

[1] In this use-case ,the data and computer systems involved have to be protected as part of HIPPA (Health Insurance Patient Privacy Act of the USA). This is an example of an influence (in this case legislative) external to the grid and to the users that put constraints on the security and possibly access management of the grid and/or grid nodes involved.

can use grid processing power and databases when required (but will be billed for this access as and when it occurs).
During the development of the service, the (now PUDS) developer invites humanities researchers to use the service for free/gratis in order to test it. Nevertheless, in his/her contract with the grid consortium, the PUDS developer has had to agree that only genuine UK academic researchers are able to use the grid resources for free/gratis, but that private individuals and 'for-profit' organisations should be billed. The PUDS developer decides to avoid the problem by only serving the UK academic community during the testing phase, but has the difficulty of checking the end-users' statuses whenever a test is made.

g) A highly technical programmer has permissions, as an academic, to use the grid. S/he writes some code and submits it to the grid to produce some computer-generated imagery (CGI) output. Once s/he receives that output, s/he is able to process it on his/her desktop machine and then re-submit it to the grid for further processing. Sometimes s/he is able to start a job running for which s/he is unconcerned whether it takes the usual two hours or three days. By prioritising his/her jobs (or by deliberately choosing the places to run them) s/he is able to use the different parts of the available grid to the greatest efficiency. (In this way s/he may be behaving as a PUA or PUS: if a mechanism were available to use slower grid resources when the priority is low, then s/he would be happy to use this and to remain a PUA. Otherwise, s/he may deliberately run high and low priority jobs in specific places, depending upon demand).

h) A satellite orbiting the earth has a grid-enabled sensor attached for use with grid research. A system administrator (GRID-SYS) has to connect to the hardware controlling the sensor to perform a firmware upgrade. This has to be done remotely and there are five individuals on the planet who are trusted to perform this task.
A GRID-SYS authenticates (somewhere), connects to the device with system administrator privileges and carries out the task.
This task needs to be carried out periodically and the five individuals change.

i) The same sensor on the same satellite is regularly switched to detect light of a different wavelength for the collaborative group of researchers across the world that uses the data. This switch must be performed manually and this can be done by about one hundred of these researchers' IT grid support staff (all GRID-SYSs).