Security, usability and the new types of grid users

A paper for xxxx

Document History

Version	Date	Comments
0.1	5 Dec 05	Absolutely work in progress. To Richard and Alun for early comments.

Contents

1. Abstract

[xxxx need to know word count] But something like

Who will be the grid users of tomorrow? Past experience looking at users of the Internet or World Wide Web tells us that from a dedicated core of highly technically gifted (in IT) individuals, other interested parties begin to take over, at least in terms of absolute numbers. Current trends also indicate that, in the not so distant future, the majority of grid users will be researchers in a variety of fields. These are likely to vary from scientists actively interested in computation to researchers needing computer power, but who are disinterested in computing *per se*. We propose a categorisation of 'future grid' users into the following categories: Service End-User, Power User (with three distinct sub-types), Service Provider and Infrastructure Sysadmin. A further basic type could be argued as Third Party Beneficiary. This paper outlines the possible characteristics of these 'types' of users. This paper will discuss briefly the levels of security, trust and responsibility that is associated with each type of user outlined above. For users that have layers of applications or, for example, a portal between them and the grid resource, it is almost certain that the heavyweight security of client digital certificates is too onerous and unnecessary. It is likely that some users will, however, need client digital certificates, due to the level of control that they could exert on grid resources. We also outline a Customer-Service model of grid use. Should the Service End-Users become the most numerous and most demanding users, we need to consider their possible requirements. This paper explores this a little and examines whether access management mediated via Shibboleth would be more appropriate for these users. It could be that authentication and authorisation for the SEU 'customers' should be the responsibility of the SPs. This hints at a more legal framework for delegating authority to enable grid use, but could be more secure and easier to administer. This model should therefore simplify the challenges of accounting on grids: leaving much of this onerous task to the Service Providers.

2. Introduction

The Pew Internet in American Life Project wrote in 2005 that 'The Web has become the "new normal" in the American way of life'. It is used by two thirds of Americans for a variety of purposes from checking email to participating in auctions. On any given day in 2004, it is estimated that 70 million American adults did something on the Internet. It was not always thus. Before Netscape's browser, Mosaic, was given away free in 1994, the Internet was the domain of the educated and technically able. Even within that educated elite, the use of the Internet was dominated by a few research subject areas, possibly arenas in which the development of computing itself had been highly relevant for many years.

Bill Dutton refs here xxxx

Grid technology must be approaching the metaphoric 'release of the browser' stage some time soon. Whether there will be a huge take-off, as was seen after 1994, or whether it will be by a more steadier take-up remains to be seen. However, it is the possibilities and ease of use to the greater community that will make the breakthrough. This paper is focussed mainly upon the educational and research use of grid technology. The engagement of the average citizen with grid technology will take much longer; however, we believe that the experience of take-up of the Internet is relevant to the divide between researchers experienced in programming or scripting and 'the rest'.

Research as to the take-up of grid technology by such 'average researchers' is difficult to find and is clearly difficult to perform. Most surveys have to obtain data from current users of grids – clearly people who have partly or fully overcome any usability issues – and therefore collectively present a very skewed picture. Anecdotal evidence of researchers refusing to engage and benefit from grid technology suggests that if an application-interface is presented that is easy to use, the uptake is strong [xxxx informal Bridges report?].

Most early adopters of the motor car were expected, and needed, to be expert mechanics. Later, the machines became more reliable and people's interest was purely in travelling from A to B: many of those same people cannot point to the carburettor. Researchers in many of the sciences during the 1960s and 70s, who were reliant upon large sets of numbers and statistics, often found it more productive to program their own spreadsheet applications if they could gain access to, or build, a computer. Later, those same researchers were able to use an 'off the shelf' spreadsheet application, and become disinterested in the technology behind the application: they could focus fully on the findings of their science.

Clearly, the more that the 'vehicles' for research – and other activities – are developed and improved, the greater the uptake is of the technology, but something else happens as well. Not only do the numbers of users increase, but the *types* of users can change drastically. Taking our reference point from the developments of the Internet and the early World Wide Web, we predict that grids will change from the enigmatic domains of highly technical computer experts to areas of greater access to all researchers. This presents a challenge to the development of grid middleware and user interfaces. Before we try to improve the experiences for the grid users of today by building up requirements, we should consider who are the grid users of tomorrow?

The following sections of this paper present our view of those users. This is a personal view, based partly on experience and partly on predictions arising from the use of the Internet and the Web.

3. Types of grid users

3.1. Categories of users and their relative abundance

Table 3.1 presents a summary of a projected set of tomorrow's grid users. In most cases, these users exist today to a certain extent and therefore their relative abundance can be compared between today and the future. There is likely to be almost a perfect inversion in terms of numbers of users in each category. This is important as much effort is expended in serving the needs of our most common type of user. If the most common type of users at present is a 'power user' or a system administrator, this can be highly distracting in terms of freeing up the power of grids to the rightfulxxxx constituency of users: the true end users. Until this is accepted, a significant barrier to entry to the main beneficiaries of the technology will remain.

Table 3.1 Grid users of the future

Type of user	Typical characteristic	Main role	Current proportion of grid users in this category	Future proportion of grid users in this category
SEU	Service End-User. Little or no computing expertise.	User of applications served by SPs.	Low	High
PUA	Power User Agnostic of grid resource node. High degree of computing expertise.	Develops programs and data but does not care where processing takes place.	Medium	Low/medium
PUS	Power User requiring Specific grid resource nodes. High degree of computing expertise.	As PUA but may have more platform etc. dependent expertise and some sysadmin expertise.	High	Low
PUDS	Power user developing a service. High degree of computing expertise.	As PUA/PUS but developing expertise like SP.	Low	Low
SP	Service Provider. High degree of computing expertise.	As PUA/PUS but has expertise in authorisation and possibly identity management.	Low/medium	Low
Grid-Sys	Infrastructure sysadmin. High degree of computing expertise.	System administration of grid nodes with possibly infrastructure delivery and security expertise.	High	Low

Note that there are clearly omissions from Table 3.1. Two notable players are the Third Party Beneficiary (TPB) and Resource Owners. A TBP could be a person or organisation who/which does not interact directly with the grid but where his/her/its personal data are being handled on the grid. Resource Owners clearly have important functions, but they do not necessarily interact with the grid, unless playing one of the six main roles shown in Table 3.1 at a particular moment in time. In designing future grids, the requirements of both of

these sets of players would have to be given much thought and would impact upon the likely architecture and security mechanisms of those grids. However, for the purposes of this paper, the general requirements (or expectations) of only the six main roles are considered in relation to access management and otherxxxx security needs.

3.2. Do we want storyline examples?xxxx

E.G. THIS KIND OF THING:

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.

3.3. Example illustrations of the six major players

Table 3.2 shows some examples of the activities and needs of these six major grid players. Where it is possible to give current real-world examples of each player's activity, this is displayed in italics. The remaining text comprises purely imaginative illustrative examples.

Table 3.2 Example activities for the six main categories of grid users

User/ player	Examples of activities
SEU	Little or no computing expertise. User of applications served by SPs.
	PhD Biologist submitting large data sets for processing to a service. A current example includes an application such as BASIS. BASIS (Biology of Ageing e-Science Integration and Simulation system, http://www.basis.ncl.ac.uk/) is a UK e-Science pilot project which delivers a grid enabled system that serves the biology of ageing research community by helping to integrate data and hypotheses from diverse biological sources. From the user's point of view the service is presented through a web portal.
	User or organisation receiving regular output (without necessarily sending input) e.g. the BBC or Meteorological Office receiving bulletins from a 'Weather' SP.
	Social scientist submitting various problems or scenarios to a social modelling and simulation service (possibly with full graphical interface as suggested by the MOSES MoSeS project (Modelling and Simulation for e-Social Science - http://www.ncess.ac.uk/research/nodes/index.shtml#moses).

PUA	Power User Agnostic of grid resource node. High degree of computing expertise. 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.
PUS	Power User requiring Specific grid resource nodes. High degree of computing expertise. 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.
PUDS	Power user developing a service. High degree of computing expertise. As PUA/PUS but developing expertise like SP. *Examples would include the developers on the BASIS project, the BRIDGES project (refs??xxxx), the NeuroGrid project (http://www.neurogrid.ac.uk/) and many more.*
SP	Service Provider. High degree of computing expertise. May have expertise in identity management and authorisation. Many of the developers, administrators and owners of projects already mentioned will play the role of SP when the applications mature. A popular method of providing this service is to build a portal, possibly using web services to give an easy interface to the SEU. XXXXAlun to assist with examples?
Grid-Sys	Infrastructure system administrator. High degree of computing expertise. A Grid-Sys may manage dedicated grid resource nodes (including clusters) and any grid system objects such as resource brokers, authentication, authorisation 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 a 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.

3.4. Access management characteristics of these players

Table 3.3 describes, blah blah…xxxx

Table 3.3 Access management/security characteristics of the six user types

User/ player	Access management/security characteristic
SEU	SEU does not need to be 'known' by a grid access management service (should one exist) as the grid trusts and accounts the SP not the user. SP may need to authenticate, authorise and account for the user as well as possibly taking on 'metering' responsibilities.
PUA	PUA need not be 'known' by a grid access management service (should one exist) but some sort of mapping to a billing account may be necessary. It is likely that a grid access management service would need status information from an identity provider for authorisation purposes.
PUS	As PUA, above. However, in addition, grid node owners may wish to have a direct authentication, authorisation (and accounting) relationship with the PUS.
PUDS	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.
SP	SP may be trusted to provide services only to those supposedly authorised to use the grid. SP may need to identify (authenticate) SEUs but will probably need to recognise status (authorisation-related attributes). SP will need strong/secure assertions of identity/authentication between it and the grid resource nodes. Accounting may be required between the grid resource nodes (or access management service) and the SP and between the SP and the SEU, although this latter requirement may not need to be met using grid middleware.
Grid-Sys	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] xxxxRETHINK THIS!!.

4. The Customer-Service grid relationship

Most common interaction on future grids. C-S etc.

Examples from BRIDGES – Richard.

5. References

Proportions of Americans on line in 2005

http://www.pewinternet.org/pdfs/PIP_Digital_Divisions_Oct_5_2005.pdf

Main reference

http://www.pewinternet.org/pdfs/Internet_Status_2005.pdf

Dutton discusses the near doubling of North American access to the Internet between Aug 95 and Mar-Apr 96, with Mosaic and Netscape introducing a graphical user interface. Dutton, W. (1999), Society on the Line (Oxford Un Press): 227-56, especially table 9.1, page 228.

Even now, the major factor shaping patterns of Internet use is skill of the user. Di Gennaro, C., and Dutton, W. H., 'Youth, Proximity to the Internet and Political Participation: The Case of Britain', Parliamentary Affairs, forthcoming.