

GRIDSITE, GACL AND SLASHGRID: GIVING GRID SECURITY TO WEB AND FILE APPLICATIONS

A. Doyle*, S.L Lloyd**, A. McNab***

* University of Glasgow; ** Queen Mary, University of London; *** University of Manchester

Key words to describe the work: Authorisation, Access Control, X509, GSI, HTTPS, GridPP

Key Objectives: To allow users of existing web browsers and file-based applications to store and retrieve data using Grid credentials.

Motivation for the work (problems addressed): More secure and less cumbersome authentication and authorisation systems are needed by websites which allow users to upload or create content. Similar security mechanisms are needed to control local file storage when applications are running across the Grid, where Grid identity is more meaningful than a local and temporarily assigned identity.

GridSite – Using Grid credentials to make a “Two Way Web.”

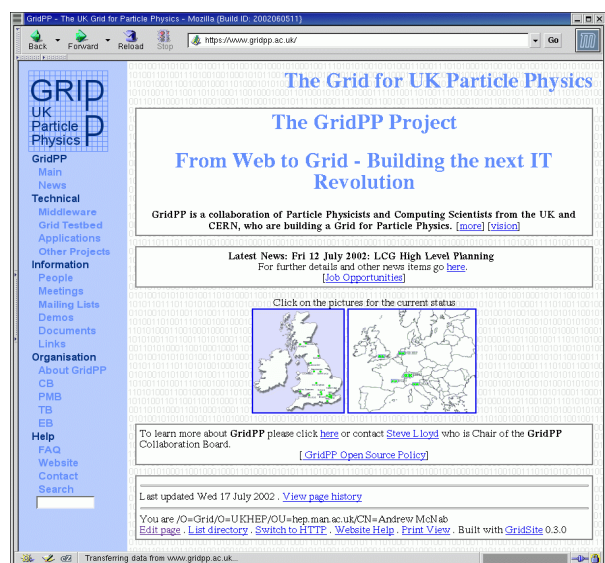
Web browsers represent the most common, familiar and most widely installed application used to access remote resources on the current Internet. However, most websites are built using HTTP technology, which can only implement cumbersome authentication and authorisation mechanisms. Typically, this involves the user choosing a short memorable password for each site to which they need to identify themselves. Consequently, the user may find themselves having to enter multiple usernames and passwords as they pass between websites run by their employer, their bank, online merchants etc. As well as the inconvenience involved, this is also vulnerable to “brute force” attacks by third parties due to the short length of the passwords.

Since the mid-1990’s, most web browsers have also supported the HTTPS protocol, which uses X509 digital certificates and has been widely used to provide authentication of websites to users. This allows a user to send credit card details to a merchant’s website, for instance, with some confidence that the site is not being impersonated by a malicious third party.

Although the corresponding user authentication to websites has been supported since the adoption of HTTPS, it has been far less used, due to the administrative overhead and cost of verifying users’ identity before giving them a meaningful X509 user certificate.

However, with the large-scale deployment of X509 certificates to members of the UK e-Science and worldwide Grid projects, this is changing, and it is now practical to base a collaboration’s website on HTTPS rather than HTTP technology, without

requiring users to install any special software.



The GridPP project, involving physicists from the UK Particle Physics community, has chosen to implement its collaboration website in this way, and to produce a general website management tool, GridSite, which is flexible enough for other projects to use for their own sites.

Since GridSite is able to uniquely and securely identify users by their X509 certificate, they can be granted rights to edit and upload webpages, images and binary files. This is enforced by Access Control Lists, using GACL, an XML schema developed by GridPP as part of its contribution to the EU DataGrid. Access control can be specified in terms of individuals or groups, with group membership managed by the group’s administrators through the same web interface.

This has allowed GridPP to devolve maintenance of the website down to the level of those directly involved in each area of work. Since the administration of group authorisation is also

devolved, the administrative overhead normally carried by the website manager is greatly reduced.

Since GridSite permits several users to maintain a set of documents, this has also made collaboration between GridPP members at different institutions considerably easier; and tools are provided to retain old versions and record document histories to automate the book-keeping of who has changed a document and at what date.

GridSite is now also in use by the EU DataGrid Testbed 1 support website, and by the UK e-Science Engineering Task Force.

As more sites using GridSite appear, the Web becomes closer to its original vision as a medium which allows users to write as well as read – a Two Way Web. A user can open their normal web browser at the start of a session and visit websites they have authorisation relationships with, modifying content, adding comments and accessing secure areas transparently, without having to manually re-establish their identify at each site boundary.

SlashGrid – Giving Grid credentials to file-based applications

Just as the Web browser is the most common way for users to access remote data, most applications use a filesystem interface to access local files on the same machine. This organises data into files, contained in a hierarchy of folders or directories, each accessible by name. For interactive use, a graphical file browser is commonly used, displaying files as icons which may be opened and accessed using a mouse. File access within an applications uses an analogous programming interface, which in most programming languages is based on a set of functions to “open”, “read”, “write” etc.

The security associated with these operations is traditionally tied to credentials which only have meaning on the machine (or in some cases the computing site or cluster) in question. Typically, this takes the form of a short username or group name, and a specific file may have one user who has permission to write to that file.

However, as we connect machines and sites together with Grid technology, these local credentials become increasingly inappropriate for managing authorisation to use resources, as they cannot readily be shared across the Grid. For

example, a user may have the username mc nab at one site, but amcnab at another, and at a third site user mc nab may be a completely different individual.

By using X509 certificates and Grid protocols we can establish a single identity across all resources we use on the Grid. GridPP has also developed extensions to the Globus software which temporarily assign a local username to a Grid identity using a computing resource. While this has been successfully used by sites participating in the EU DataGrid Testbed 1, the system cannot readily be used when creating long lived files, since the username they are “owned” by is only temporarily associated with a specific Grid identity.

To resolve this shortcoming, we have produced a file system framework, SlashGrid, which allows file and directory authorisation to depend on long-lived Grid identities. Initially produced for the Linux operating system, SlashGrid creates a hierarchy of directories under /grid where an application’s username, whether static or temporary, is irrelevant to whether it can create, read or modify files – what matters are the Grid credentials the application currently holds on behalf of the user, wherever they are on the Grid.

For interoperability with other products of the EU DataGrid and related projects, SlashGrid uses the same GACL toolkit as GridSite, and the same Grid Access Control List XML schema.

SlashGrid has also been designed to be readily extensible, by the use of third-party plugins to add additional filesystem types. In particular, we have implemented an HTTP filesystem, in which the contents of remote websites can be accessed by applications as if they were local files.

When combined with the secure authentication of protocols like HTTPS and GridFTP, this has the potential to allow existing applications to operate on the Grid, indifferent to the true location of the files they manipulate, with remote Grid file access provided as a service by the operating system layer.

More information about the Open Source products GridSite, GACL and SlashGrid can be found at <http://www.gridpp.ac.uk/projects.html>