Audited credential delegation: a usable security solution for the virtual physiological human toolkit

Ali N. Haidar¹, Stefan J. Zasada¹, Peter V. Coveney¹,*, Ali E. Abdallah², Bruce Beckles³ and Mike A. S. Jones⁴

¹Centre for Computational Science, University College London, 20 Gordon Street, London WC1H 0AJ, UK
²E-Security Group, London South Bank University, 103 Borough Road, London SE1 0AA, UK
³University of Cambridge Computing Service, Pembroke Street, Cambridge CB2 3QH, UK
⁴Research Computing Services, Devonshire House, Precinct Centre, The University of Manchester, Manchester M13 9PL, UK

We present applications of audited credential delegation (ACD), a usable security solution for authentication, authorization and auditing in distributed virtual physiological human (VPH) project environments that removes the use of digital certificates from end-users’ experience. Current security solutions are based on public key infrastructure (PKI). While PKI offers strong security for VPH projects, it suffers from serious usability shortcomings in terms of end-user acquisition and management of credentials which deter scientists from exploiting distributed VPH environments. By contrast, ACD supports the use of local credentials. Currently, a local ACD username–password combination can be used to access grid-based resources while Shibboleth support is underway. Moreover, ACD provides seamless and secure access to shared patient data, tools and infrastructure, thus supporting the provision of personalized medicine for patients, scientists and clinicians participating in e-health projects from a local to the widest international scale.

Keywords: grid security; e-health security; information assurance; security wrappers

1. INTRODUCTION

Within the virtual physiological human (VPH) initiative (www.vph-noe.eu), grid infrastructure provides access to a wide range of computing resources distributed across multiple administrative domains. Scientists and clinicians need to use such resources to perform patient-specific modelling and simulation that draws on the medical characteristics of an individual patient. Decision-support systems based on patient-specific computer simulation hold the potential to revolutionize the way clinicians plan courses of treatment for patients [1]. This leads immediately to the question of how to address information security within the VPH initiative.

As high profile security breaches and data loss are frequent headline news [2,3], a usable security solution is of critical importance for VPH projects. There are several pieces of legislation such as the UK Data Protection Act, the EU Data Protection Directive and the US Health Insurance Portability and Accountability Act (HIPAA) that make it a legal requirement for VPH partners to collect, hold and process patient data in a secure way [4]. Security is also needed to protect VPH projects from the consequences of unauthorized disclosure of medical information including negative publicity, legal liabilities and fines; and from unauthorized modification of patient data used in VPH project environments, which may lead to incorrect patient treatment and result in a loss of life or identity theft, itself currently creating considerable concern. Hence, authentication, authorization and auditing security mechanisms are key requirements for any VPH system using patient data to be compliant with information security standards and avoid legal liability.

Another major problem faced by end-users and administrators of grid-based VPH environments arises in connection with the usability of the security mechanisms deployed [5]. Many of the existing computational grid security infrastructures use public key infrastructure (PKI) and X.509 digital certificates as the means to provide authentication and authorization security goals. For instance, Globus (www.globus.org), UNICORE (www.unicore.eu), virtual organization membership service (VOMS) [6] and community authorization service [7] are all based on PKI [8]. However, it is well
documented that such security solutions lack user-friendliness [5,9] for both administrators and end-users, which is essential for the uptake of any VPH solution. The problems stem from the process of acquiring X.509 digital certificates, which can be a lengthy one including the generation of proxy certificates to get access to remote resources as part of the authentication process (see the electronic supplementary material, §1). As a result, many users engage in practices which substantially weaken the security of the environment, such as the sharing of the private key of a single personal certificate, to get on with their tasks.

End-users, such as scientists or clinicians who are not security experts, are concerned with the results of the analysis they perform on such grids rather than acquiring and using digital certificates [5]. Administrators are concerned with setting up virtual organizations (VOs) and administering security infrastructure in an efficient way. Resource providers are concerned with securing access to their shared resources, tracing users responsible for performing tasks on their resources, and avoiding the consequences of security breaches, including negative publicity and fines. Moreover, there is a need within the VPH initiative for a security solution that can be easily integrated with the tools provided by the VPH Toolkit [10]. These software tools have been developed by various partners and third parties using different programming languages to access and process patient medical data. Without such security, each set of VPH tools would need to have a ‘hard wired’ security extension in order to be compliant with data security standards. This also means that VPH users would have to maintain credentials for all these VPH tools, which would be difficult to manage and would probably deter clinical uptake of VPH approaches.

This paper describes the application of the audited credential delegation (ACD) [11,12] security solution to address authentication, authorization and auditing security goals within grid-related projects, including VPH and many other projects. We show how ACD satisfies security and usability. We demonstrate how ACD can be used to set up multiple VOs that have specific goals within the VPH initiative, to manage dynamic groups of users wishing to access various resources, and to provide VO administrators with tighter control of users’ actions as well as identity management. ACD is more than simply a security layer. Existing solutions such as MyProxy, Shibboleth and SARoNGS only provide credential repositories to store short-lived X.509 certificates (Myproxy), web-based single sign-on (Shibboleth), and web portals to access grid resources using a combination of Shibboleth and VOMS (SARoNGS) [9,13]. None of these solutions provides a holistic VO-controlled security solution in the way ACD does.

We have successfully integrated ACD with the functionality of the application hosting environment (AHE) [14], lightweight grid middleware that allows the user to run applications on the grid, to construct a VO with tight security controls on identities and actions while providing a set of services allowing users to interact with grid resources without requiring specific knowledge of the details of each resource they wish to use. In addition, we have integrated authentication, authorization and basic auditing of ACD with the Individualized MEdiciNe Simulation Environment (IMENSE) [15], developed within the ContraCancrum Project (www.contracancrum.eu), to provide secure access to clinical data and tools. The functionality used in the environment includes the performance of imaging data annotation and analysis, the running of simulations and composite tasks (workflows) of considerable complexity on remote grid resources using patient data. The integration of IMENSE with ACD provides assurance about the confidentiality and integrity of patient data because only authorized scientists and clinicians are able to view and modify patients’ clinical records as well as having easy and controlled access to remote grid resources using familiar authentication mechanisms.

The paper is organized as follows. Section 2 gives a brief overview of the current security challenges encountered within VPH, namely enabling scientists to access grid infrastructures and providing secure access to shared patient data. Section 3 provides a brief overview of common VPH projects’ security requirements. Section 4 presents a description of ACD. Sections 5 and 6 describe two case studies which demonstrate how ACD can be integrated with VPH environments to enable secure and usable access to patient data and grid infrastructures. Section 7 discusses related work, while §8 contains a discussion and conclusions.

2. OVERVIEW OF CURRENT SECURITY ISSUES WITHIN VIRTUAL PHYSIOLOGICAL HUMAN

This section describes two major security issues encountered in VPH environments. The first concerns the complexity of current mechanisms for accessing grid resources; the second addresses secure access to shared patient data for VPH collaborators.

2.1. Access to grid resources

To illustrate the complexity of current mechanisms for accessing grid resources, such as those provided by the UK National Grid Service (NGS) (www.ngs.ac.uk), US TeraGrid (www.teragrid.org) and EU DEISA (www.deisa.eu), we briefly describe the current steps needed by a scientist prior to running any application on a grid. For more details, the reader is referred to Haidar et al. [16] and the electronic supplementary material. The first step is to acquire a digital certificate. There are three processes involved in this step, each of which has a mean duration of one working day. The certificate authority (CA) informs the registration authority (RA) that a user has applied for a certificate (1 day). The RA contacts the user and arranges a face-to-face visit (1 day); the CA then issues the certificate (1 day). The average scenario takes about three days, which is too long. In the second step, the user is required to get authorization to access the resources offered by the resource provider. From our experience, this step takes between 3 working days and two weeks but only needs to be performed once. In the final step, end-users have to configure their chosen client applications themselves, including the Globus toolkit, the UNICORE Client and the AHE client.
which are used to access the grid with a certificate. The resource providers patently cannot do this because they have no control over or access to the end-users’ machines. An exception would be where the user invokes a web portal. All in all, the above steps amount to a lengthy and complicated process which certainly deters many potential VPH users from exploiting the enormous power locked up within grid resources.

2.2. Secure access to shared patient data

Currently, scientists working within VPH projects collect pseudonymized or anonymized patient data from hospitals (this may include patient records, histopathological and molecular data, magnetic resonance imaging, X-ray computed tomography and positron emission tomography imaging data) and upload them to their VPH environments. These data can be stored in a centralized data warehouse or distributed across several administrative domains. When the data reside within an environment managed by a VPH research group, it is by no means clear what security measures are taken to protect these data. Recent studies [5,17] have shown that many VPH and other e-Science projects do not have adequate security solutions in place to protect patient data. Although patient data are anonymized or pseudonymized by the providing hospital, it can still conceivably be identified in various cases. For example, genetic sequence data taken from a person at an interview, whose identity is therefore known, could be compared with anonymized data stored in a database; if a match were found that person’s medical status would then be revealed. An incident reported in 2008 [18], where a nurse’s medical status was revealed publicly in an unauthorized way by a colleague in the hospital where she worked, illustrates the impact of such breaches of confidentiality. The nurse’s medical history showed that she had been treated for HIV. The revelation resulted in her contract not being renewed by the hospital and her colleagues at work knowing about her disease. The hospital was ordered to pay the nurse €14,000 in damages and €20,000 in costs.

Therefore, there is a very obvious need for a secure solution that enables VO-controlled access to patient data within VPH projects to ensure patient confidentiality and integrity, along with secure and seamless access to remote grid resources for processing such data.

3. COMMON SECURITY REQUIREMENTS IN A VIRTUAL PHYSIOLOGICAL HUMAN ENVIRONMENT

In order to design a usable solution to access grid resources and patient data within VPH projects, it is fundamental to understand all the stakeholders’ requirements. The stakeholders in VPH environments include patients, scientists, clinical researchers and clinical practitioners, system administrators, universities, and grid resources providers. Scientists and clinicians need to:

— query patient data and access data analysis tools;
— invoke familiar and usable security mechanisms to perform their tasks; these must not be a barrier to their progress, and so must be seamlessly integrated with their desired ways of working.

System administrators require a mechanism for setting up VOs and administering the VPH security environment in a clear and easy fashion. This requires understanding of:

— how a scientist from a VPH project becomes a VPH user with access to grid resources;
— how to authenticate VPH users to resource providers; and whether VPH users can use their local credentials (preferably the same ones they use in their own organization) to access grid resources or need to acquire new ones;
— how to determine whether a person within a VPH project is authorized to perform a task on a grid resource;
— who decides what the access rights of a VPH scientist are;
— how to identify those people from VPH environments responsible for performing tasks on grid resources using patient data.

Resource providers, in particular the hospitals providing patient data together with grid resource owners, are concerned with securing access to their resources. This involves identifying who is requesting access to their resources (authentication), checking if a user is allowed to run tasks on their resources (authorization) and tracking users responsible for running named tasks on their resources (auditing) in case of misuse (e.g. security breaches, usage of CPU allocations for billing purposes). All these measures are needed to give resource providers assurance that their assets are adequately protected and to ensure that the resource providers avoid the consequences of the misuse of their valuable resources by unauthorized users.

4. AUDITED CREDENTIAL DELEGATION

4.1. Overview

The design of ACD is based on the concept of ‘wrappers’. A wrapper is a connector between a component and the outside world. It enables controlled access to the functionalities of a component. For instance, figure 1 shows the ACD security wrapper made of authentication, authorization and auditing components surrounding the functionalities of an environment represented by the tasks (Task1, …, Taskn) that can be performed on the system. Any request by a user to perform a task is intercepted by each layer of the security wrapper to establish the identity of the requester, to check whether or not the user is allowed to perform the task, to record the results of these checks in the audit log, then to perform the task on the system and, finally, to return results to the user.

This model fits well with many VPH environments that encapsulate tools from the VPH Toolkit [10] as
we will show in §§5 and 6. These tools are usually specified as ‘black boxes’ so that scientists can use them to access patient data without knowing their internal details. The interface of the tool is the only information available to the designer about how it will be connected with its environment. These tools have to be customized in some way to match the global requirements of the VPH environment described in §3, such as the need for extra security features or blocking unneeded functionality provided by a tool. By placing VPH tools within a security wrapper such as ACD, all the requests coming to and/or replies from the wrapped tools are passed through the authentication, authorization and auditing wrappers. These security wrappers hide the details of the interface of the tool from external clients and act as an interface between its caller and the wrapped tool. The interface of the wrapped tool is different from the interface of the security wrapper. The wrapper’s interface will include the names of the tasks provided by the wrapped component in addition to the tasks provided by the security wrapper. The security wrappers will define how a call to perform a task offered by the wrapped component will be processed. In this way, ACD controls who can access the specific functionality provided by a VPH tool, determines whether the user is allowed to access the functionality and traces users who have invoked this functionality. Without such wrappers, the interface of a tool is accessed directly without any protection.

ACD provides much of the functionality required for secure cloud computing [19], a business model of grid computing, that provides access to various resources such as CPU, memory and storage (known as infrastructure services) and applications. However, it is not designed to be a cloud computing security solution. Amazon’s Elastic Compute Cloud (EC2) and Google App Engine are examples of such clouds [20]. There are many security issues in cloud computing that are yet to be resolved concerning data storage, compliance of the cloud system with legislation (DPA, HIPPA) and information assurance [19,20]. The main difference between clouds and VOs used in ACD is that the VO has full control of where data are stored and the processes that access these data whereas within the VO, in a cloud environment, the service and data maintenance are provided by third party vendors, potentially leaving the client ignorant of where the processes are running or even where the data reside. The location of data storage is very important so that applicable laws and regulations governing the data are identified [4]. Only recently, Amazon and Microsoft started offering data storage guaranteed to be in Europe to address the legal aspect. Users of cloud services have to trust the provider as to where and how the data are protected and the adequacy of the security controls in place, both critical issues for VPH projects.

The design of ACD has been focused around several objectives. First and foremost is the requirement to provide secure yet facile access to grid resources and to ensure the confidentiality and integrity of patient data used in a research environment. There is a need for a solution that can be easily extended, because new tools are developed during the lifetime of VPH projects as well as acquired from third parties; these also need to be exposed to end-users in a secure way. Keeping this in mind, ACD has been designed around Web services, providing interfaces compliant with Web services standards such as web service description language, SOAP, WS-Policy and WS-Security [21]. This enables integration of new VPH tools written in programming languages that have Web services libraries with ACD. In addition, ACD has been developed by adopting best
practice software engineering principles that enable it to evolve as new functionalities are needed or changes in security policies are required, without the need to rewrite the whole solution from scratch or perform major modifications. Besides secure access to patient data, ACD enables VPH scientists to seamlessly access grid resources using various authentication mechanisms such as a local ACD username–password, or Shibboleth credentials, both of which are considered easier than acquiring and managing digital certificates, in order to run pre-installed applications on AHE, such as complex workflows and simulations that support patient-specific treatments. By providing support for Shibboleth, a large class of end-users who belong to institutions subscribed to Shibboleth services (e.g. academic institutions) will be able to invoke their local institutional credentials rather than acquiring a VO specific username–password. Within VPH, the correct execution of ACD functionalities to ensure integrity and confidentiality of patient data is extremely important. Hence, at the outset of its design, ACD was subjected to a rigorous modelling activity based on formal methods to ensure that the security requirements were fully met [12].

Another critical aspect addressed during the design of ACD is usability. ACD eliminates the steps performed by end-users listed in §2.1 which are now done only once by an expert-user (the VO administrator). It is important to emphasize that the time consuming steps described in §2.2 cannot be completely eliminated because of the need to interoperate with grid resource providers’ systems. What we have improved is that if there are say 10 scientists in a group, only one person (the expert user) has to go through the steps whereas the others will enjoy genuinely seamless access thereafter. Hiding complexity from end-users whenever possible is a fundamental usability principle. We do not claim that there are no usability problems with passwords but the usability issues associated with digital certificates are substantially worse. A digital certificate used to access grid resources is supposed to be protected by a passphrase (i.e. a password), so with digital certificates we still have all the usability problems associated with passwords as well. We have recently completed a comprehensive usability study [22] that involved comparing several middleware products for accessing grid environments. These include the AHE middleware, introduced in §1 and described in detail in §5.1, which comes with graphical user and command line interfaces for accessing grid resources, a combination of AHE with ACD, as well as UNICORE and Globus. There were 40 participants drawn from different departments and faculties at UCL including Physics, Chemistry, Computer Science, the Medical School, the Business School, the Cancer Institute and the Law School. Each participant was asked to run a simulation on a grid (NGS) using the different middleware to configure the security of their client tools and use the credentials given to them (username/password, X.509 certificate). The results unambiguously show that the combination of AHE and ACD scored higher than all other tools regarding the time needed to run a task, the ease of configuring the security of the tools, and the ease of running the overall task.

4.2. Overview of ACD Architecture

ACD has four components:

— A local authentication service (LAS): one of the main objectives of ACD is to remove digital certificates from the end-users’ experience. The current implementation supports a username–password database specifically for ACD. To be authenticated, a user has to provide a username–password pair that matches an entry in the database. To avoid known vulnerabilities in usernames and passwords we adopted OWASP best security practices [23] such as storing passwords in an encrypted form, rejecting weak passwords chosen by users, forcing the password length to a minimum of eight characters including special characters, and changing the password on a regular basis. This way, if the database is compromised, the attacker will not get hold of any password. There is currently work in progress to support Shibboleth in ACD to give users more options to choose from. Shibboleth is currently used by many universities in the UK and EU to allow students and researchers to access online publishers’ resources by invoking their local university username–password credentials. This way they will not need to use a specific ACD username–password for the VO. However, the support of Shibboleth will have an impact on ACD availability since it is dependent on the availability of the external authentication services provided. Without successful authentication, it is not possible to determine the role of the user in any given VPH project and, as a result, all requests to perform tasks will be denied.

— An authorization component: this component controls all actions performed in the VO. It uses the parameterized role-based access control (PRBAC) model in which permissions are assigned to roles [24] as shown in figure 2 (Role → Task). The VO policy designer associates each user in the VO with the role that best describes his/her job functions (UserID → Role). The policy is defined at the VO set-up because it depends on the VO functionalities. The tasks (permissions) assigned to roles are drawn from the VO functionality. Sections 5 and 6 show how this is done. There are administrative tasks common to all VOs, such as ‘create role’, ‘assign a VO user to one or more role’, ‘assign tasks to roles’ and so on. This component is usually configured during the VO set-up by the VO administrator. In traditional role-based access control, two users that perform similar roles in the VO must have identical permissions. Sometimes this is not desirable. For instance, when two scientists submit two jobs to a grid resource, each scientist should be able to privately monitor, terminate or view the result of his/her own job submission. Thus the PRBAC model is flexible and permits fine-grained access control. It is important to emphasize that the decision to permit a user to perform a task on a grid resource is determined by the resource provider who has the final authority. The VO authorization component only manages the permissions (i.e. the allowed tasks) given by the resource owner to
the VO which controls the use of these permissions within the VO (authorization delegation).

— A credential repository: this component is responsible for managing the delegation of identity from the user to ACD via a proxy certificate. It stores the certificates acquired by the VO administrator through the steps in §2.1 and their corresponding private keys in order to communicate with the grid (Certificate → Key). The relation ProjectName → UserID enables the creation and management of VO membership. In order to allow the members of a named VO access to grid resources, the VO is assigned a digital certificate (ProjectName → Certificate) which is used behind the scenes to authenticate requests issued by the VO at the resource provider site. The component also maintains a list of issued proxy certificates (delegated identities), their corresponding private keys (Proxy → Key) and the association between users and proxies (Proxy → UserID) in order to trace which proxy was used by which user. These proxies enable users’ requests to be authenticated at remote grid resources (known as identity federation) on behalf of the users. At the grid resource owner’s end, all requests to access grid resources appear as coming from the named VO, not individual users. Two users who submitted jobs on the same grid resource site will have different proxies issued by the same VO certificate. The resource provider will not be able to tell which individual used this proxy to run an application on its resources but ACD can provide this information. The grid resource owner provides the VO administrator with the proxy’s public key. From the relation (Proxy → UserID) the VO administrator can tell which person used this proxy and take any appropriate action.

An auditing component: this component records all actions within the VO including authorized and unauthorized requests to perform tasks within the VO, the username that requested them, the number of login attempts and login times. This allows the VO management to identify those ACD users responsible for having performed any tasks in a VPH environment.

The main features of this architecture are the identity delegation and authorization delegation which are handled by a trusted entity, the VO, to make access to remote grid resources easier and to provide finer access control decisions within the VO. Since end-users sometimes share certificates to get access to shared resources, ACD is just an organized way of doing so thereby mitigating and controlling the risks associated.

5. INTEGRATION WITH THE APPLICATION HOSTING ENVIRONMENT

This section describes how ACD is integrated with the AHE to enable construction of VOs that enable scientists to run pre-configured applications on remote grid resources using ACD username–password credentials.

5.1. Overview of application hosting environment

The AHE [14,25] is a lightweight mechanism for exposing scientific applications (i.e. workflows and complex simulations) as Web services, and allowing users to interact with those applications using simple client tools (AHE client). AHE enables the launching of pre-existing scientific applications installed by an expert.
user on a variety of different computational resources, from national and international grids of supercomputers, through institutional and departmental clusters, to single processor desktop machines [26]. The end user is presented with a choice of very lightweight clients, specifically designed to obviate the need to deal with Globus and UNICORE middleware for job management, allowing the user to submit, monitor and download application results, as well as to terminate applications as they run.

5.2. AHE with ACD: usable and secure access to grid resources

The current security model for AHE requires each individual VPH user to have a digital certificate, which carries with it the need to go through the steps described in §2.1. In order to remove the need for such a certificate, we have integrated ACD with AHE. The first step of the integration requires understanding the interface of AHE and ACD combined, in other words, the functional and administrative tasks that can be performed within the integrated system. The administrative tasks offered by ACD include create VO, assign certificate to VO, add user to VO, reset user password, create role, assign tasks to roles, and assign users to roles. The functional tasks offered by AHE include prepare job, submit job, monitor job, download and terminate job. Note that AHE’s functional tasks are the same as the tasks permitted for any authorized user on a grid resource site that uses Globus or UNICORE middleware such as in NGS, DEISA and TeraGrid. Therefore, the permission assignment to the VO is done by the grid resource owner first, then the VO administrator re-assigns these permissions to the roles in the VO according to the VO authorization requirements.

In the combined ACD + AHE environment, the authorization requirements determined by the VO administrator are expressed through the introduction of two roles: VO administrator and scientist. The former is permitted to perform all the administrative operations above in addition to terminate, monitor and download any job submitted to grid resources. The latter is permitted to perform all AHE operations in such a way that a person who submitted a specified job can only perform AHE functional operations on this application. As a result, two VPH users running applications using different patient data will not be able to view the results of each other’s digital activities. In addition, the scientist role only permits a user to change his/her own password.

The construction of a VO requires that an expert user goes through the lengthy process described in §2.1. Once this is done, the VO administrator creates a VO (see supplementary document §2) and assigns the certificate to the named VO using the AHE + ACD client. Then, it becomes possible to add users instantly to the VO and give them genuinely seamless access to grid resources. To illustrate how this system works consider a user named ‘John Smith’ who is a member of a research group in a UK university and would like to use NGS grid resources to run scientific applications using AHE. The user contacts the local VO administrator and requests an account. The VO administrator creates a new user account which generates a username and a random password that are given to the user. The VO administrator assigns the user to the ‘scientist’ role described above and assigns the user to a VO that has access to NGS resources (figure 3). When a user logs in for the first time to the AHE + ACD client application, he is prompted to change his password. The communications between the AHE + ACD client and the wrapped AHE server, as well as between the latter and the grid resources, are protected by the SSL security protocol.

In order to submit a job to a grid resource, the user invokes a request to perform the ‘submit job’ task within the combined AHE + ACD client as shown in figure 3 (1). This request is intercepted by the ACD authentication component which checks whether the username and password match an entry in the database. The result of the authentication is recorded in the auditing component (2). The role of the user is picked up from the authorization component, userID → [Role], in this case ‘scientist’. The authorization checks whether the task ‘submit job’ is permitted for the ‘scientist’ role held by the user, which is true (3). The result of the access control check is recorded in the audit log (4), and the operation ‘submit job’ is invoked from the AHE server (5). Once the request is granted, ACD picks the certificate associated with the VO the user wants to use (i.e., NGS) and checks whether the user is assigned to this VO. If the check is successful, then ACD generates a proxy certificate from the VO-assigned certificate, ProjectName → Certificate (6), uploads it to the MyProxy server (7) and records the issued proxies, Proxy → UserID (credential delegation occurs here), in the credential repository. ACD sends the randomly generated username/password pair needed to access MyProxy to the AHE server to download the session proxy (8) and (9). Finally, the AHE server sends the request to the grid resource site along with the proxy. At the NGS site, the proxy is validated, since the proxy is issued from a valid trusted certification authority. Certificate authentication succeeds, and the distinguished name on the proxy (VOName) is checked against the grid-map file within the NGS authorization system to determine the role of the VOName, which is Scientist. Since this role is allowed to submit a job to NGS the task will be invoked. From NGS’s perspective, it is the VOName that submitted the task, not ‘John Smith’. In order to find out who invoked the ‘submit job’ task on NGS using a specific proxy, the NGS administrator passes the public key of the proxy to the VO administrator who can identify the name of the user from (6), which records the issued proxy in Proxy → UserID. In this way, requests from within the combined ACD + AHE are audited. It is thus possible to identify legitimate users and to ensure that only such users are allowed access to grid resources, in conformance with the policies enforced by the grid infrastructure management. In addition, it is possible to detect unauthorized attempts to access resources from within the VO and to identify persons responsible for such attempts. This form of accountability is an essential requirement for resource providers to be prepared to accept the ACD security model.
To illustrate how unauthorized requests to access resources are detected, let us assume that the above user is attempting to invoke the ‘remove user from a VO’ task, which is only permitted to a user holding the role ‘administrator’. When the request reaches the authorization wrapper in (3), the current user’s role is determined, which is ‘scientist’ and it will not find the requested task among the permitted tasks for this role. As a result, the authorization wrapper will return ‘access denied’ and record this result in the audit log (4). After three unauthorized access attempts, the VO administrator is notified by email via ACD that the user named ‘John Smith’ has had three unauthorized attempts to perform the task ‘remove user from a VO’ task. The VO administrator can then take the appropriate action.

6. INTEGRATING AUDITED CREDENTIAL DELEGATION WITH THE INDIVIDUALIZED MEDICINE SIMULATION ENVIRONMENT

6.1. Overview of IMENSE environment

One of the main objectives of the VPH ContraCancrum (Clinically oriented translational cancer multilevel modelling) project (www.contracancrum.eu) is to provide an environment (it can also be thought of as a VO) that allows clinicians and researchers to use the tools developed as part of their clinical and research practice in order to run workflows and simulations on grid infrastructure, using a heterogeneous set of patient data provided by the University of Saarland Hospital within an integrated IT environment, known as Individualized MEdiciNe Simulation Environment (IMENSE) [15]. These data include heterogeneous image scans (i.e. MRI, PET, CT), patient records, histopathology data and DNA profiles. The main functionalities provided by this VO include the ability to bring together and query patient data, edit them, upload and download image data, and to invoke Web services that allow workflows including simulations to be run on grid infrastructure. For example, a workflow that checks whether a patient responds to a particular drug is a pre-configured application in AHE. For the end-user, the workflow is viewed as a ‘black box’ and users can only run the workflow using a specific patient dataset and download the results (see §3 in the electronic supplementary material). ACD only controls access to the interface of the workflow. We use DEISA and TeraGrid for large-scale computationally intensive patient-specific workflows that involve moving data from within the VO via an un-trusted public network to remote grid
resources. Thus, the following security requirements need to be addressed:

— restricting access to the environment to authorized users only;
— enabling members of the project to run applications on grid infrastructure using username and password only;
— allowing users responsible for running a given task on the environment to be traced;
— ensuring the integrity of patient data by controlling the tasks that process these data in order to offer medical treatment;
— protecting patient data when transferred onto public networks.

Prior to the integration, access to IMENSE functionalities did not meet the above requirements.

6.2. Integration of ACD with IMENSE environment

Having understood the functionalities of IMENSE introduced in the previous section, the integration with ACD can be done as follows. The administrative operations of ACD remain as described in the previous section. However, the functional activities performed within IMENSE now include uploading and downloading patient-specific images, running workflows on patient data, viewing images, searching patient data and image segmentation inter alia. The authorization requirements for this system are expressed again through the introduction of two roles: VO administrator and scientist. The first role is permitted to perform all the operations above. The ‘scientist’ role is permitted to perform all the functional operations, in addition to enabling the user holding this role to change his/her own password. The result of the integration is a controlled VO within which each request to perform a task goes through all three security wrappers previously described: authentication, authorization and auditing. We illustrate this through an example (see figure 4).

A user can join the IMENSE VO in the way described in the previous section. Consider the same user ‘John Smith’ who wishes to run image segmentation on appropriate grid resources. The request to perform this task is first intercepted by the authentication wrapper which checks the user credentials against the ACD authentication service. The outcome of the authentication is recorded in the audit log. After successful authentication, the role of the user is determined from the authorization component (userID → [Role]), which is ‘scientist’, permitted to perform the ‘image segmentation’ task. The result of the access control check is also recorded in the audit log. Once access is granted the task is performed in the VO; as a result, all the steps described in the previous section steps (1) to (11) needed to run ‘submit job’ are performed behind the scenes to run the image segmentation application pre-installed on AHE. Once segmentation finishes, the user is notified to download the result. The same level of auditing is also provided in this environment. This ensures that only authorized personnel can run tasks in the VO and that the user can only access the result of the segmentation request they submitted. The permissions in the VO are assigned to roles by the VO policy designer who understands what the users require in order to do their jobs.

7. RELATED WORK

There are certainly precedents for the concept of VO's used in ACD whereby users invoke either their local credentials or a dedicated username and password, such as in the ‘community account’ system provided by TerraGrid [28] and SARoNGS [13] offered by NGS. For instance, the community account system allows scientists to access grid resources using a dedicated username and password via a Web portal. The SARoNGS project shares various similarities with our approach. It removes digital certificates from the end-users’ environment, enabling them to invoke their local credentials via a Shibboleth federated identity system, which is then translated into a grid identity credential to access UK NGS grid resources. It differs from ACD in that it passes individual identity and attributes of the user to the grid layer whereas ACD presents a single identity (that of the ACD VO name). The SARoNGS approach assumes the use of a web portal and requires an end-user (or portal on behalf of the end-user) to specify VO membership and role parameters before being able to access the grid. Like ACD, the mechanism is based on providing easy access to grid resources. The main difference is that ACD controls the authorization decision for the VO, whereas SARoNGS merely propagates authentic information about users and their roles within their specified VOs to the resources where it is consumed and processed. Thus, a significant part of the authorization in SARoNGS takes place within the grid resource provider's service whereas ACD assumes the role of a delegated authorization decision maker for those resources. The SARoNGS model is essentially the VOMS model [6] with Shibboleth presented to the user and the grid X.509 Certificates hidden [13]. The advantages of ACD over SARoNGS are that the VO members’ activities can be more tightly controlled (helping VO-based security) and managed (delegating responsibility for usability to the VO and the AHE). A limitation is that resource providers can only make their authorization decisions on a VO level: they are not be able to identify individuals without consulting the ACD VO administrator.

It is important to emphasize that what we present in this approach is a holistic VO-based authorization solution which has control of actions as well as identity. This is not the case in any other established grid environment. We have integrated our work with an environment which allows the user to actually run applications on the grid (namely the AHE); ACD is not simply a security layer, as in MyProxy, Kerberos, Active Directory, Shibboleth or Fermilab’s security mechanisms [9]. These security components only address authentication issues whereas ACD addresses authorization and accountability as well. Some of the comparisons between the examples cited above and ACD are discussed in Beckles et al. [9]. The Member
Integrated X.509 PKI Credential Services (MICS) (http://www.tagpma.org/authn_profiles) is a profile used in technologies such as MyProxy CAs. These, however, focus on providing the user with certificate-based credentials for authentication, do not deal with VO/Community attributes and leave authorization to the resources alone; by contrast, ACD in combination with AHE manages VO-specific authentication and authorization. Any solution which involves each end-user having to obtain an individual certificate (even if they immediately deposit it in a credential repository and thereafter employ a username and password to access the certificate in the repository) is unsuitable because the end user will still have to go through the steps described in §2.1.

CROWN [29] and gLite [30] middleware adopt the Globus security model and use X.509 certificates for authentication, one of the main problems ACD solves. gLite also uses the VOMS model for authorization. Unlike CROWN and gLite, authorization in ACD has been extended to the end users’ technical environment to provide fine-grained access control. This fits naturally within the VO model because, from a remote resource provider’s perspective, all VO users appear as a single user since the VO certificate is used to generate the proxies on the users’ behalf. In all the above alternative security solutions, auditing is performed at resource providers’ sites. In case of a security breach, the VO management relies exclusively on the individual resource provider’s audit logs. ACD provides auditing for every VO set up based on the tasks that need to be monitored. These tasks are derived from the functionality of the VO and, moreover, allow VO management to corroborate resource providers’ claims in case of a security breach.

8. DISCUSSION AND CONCLUSION

The ACD security mechanism has required an evolution of grid security policies because it violates the standard one-user-one-certificate security model prevalent in current grid infrastructures. A key requirement from resource providers in order for them to consider the ACD security model is the ability for them to audit all actions related to accessing their resources. This is addressed by the fine-grained auditing features of ACD. The combined ACD + AHE is now listed among the gateways on the TeraGrid Science gateways (www.teragrid.org/web/science-gateways/gateway_list) that are allowed to provide a community of users access to TeraGrid resources using the ACD security model.

ACD integrated with AHE has been successfully deployed on TeraGrid, NGS and DEISA. A detailed usability study involving undergraduates, scientists and system administrators will be published in the near future [22]. A small-scale pilot usability trial of
this security architecture, in which it is compared with the traditional PKI-based authentication mechanisms used in many existing computational grid environments, has already shown that users favour the familiar username--password paradigm supported by ACD. While that study only involved undergraduates at UCL with no prior experience of using computational grid environments, the findings are fully borne out by the extended study [22]. Usability issues associated with username–password combinations remain but they are easier to deal with than those of digital certificates.

ACD addresses many common security requirements such as the one described in §3. However, some projects that deal with data that can identify individual patients might require a higher level of assurance (LoA), meaning that the username–password dual on its own might not always be sufficient. ACD supports the National Institute of Standards and Technology (NIST) [31] LoA level 1 at best because there is little control of where a GSI-Proxy credential is kept, how it is protected, its cryptographic makeup, and its longevity. Certainly this could be improved but ACD’s main focus is user management and controlled access first and foremost and not about upgrading the entire infrastructure to cope with multiple (higher) LoAs. The LoA required will depend on the sensitivity of the shared data. This requires a vulnerability assessment of the various types of patient data (e.g. MRI and PET scans, genetic sequences) that describes the impact of loss of data confidentiality, integrity and availability so that appropriate security mechanisms can be deployed. Once these vulnerabilities are understood, it is possible to choose the appropriate security control to mitigate the risks. For instance, there might be a need for using two level authentication that involves a pin number in addition to a username–password pair, as currently employed in online banking security systems.

ACD balances different risks. On one hand, the ACD delegated authentication model may lead to the situation wherein one misuse may result in the whole VO being blocked; it is therefore essential to the VO that it vets and controls activities because the scale of withdrawal of service is much more of an issue than for an individual user. On the other hand, an individual should be encouraged by the easy access to grid resources and therefore very likely make far greater use of these resources.

ACD fits well with the distributed computing requirements of the VPH initiative and translational, computationally based biomedical research more generally. A dedicated VO for clinicians and scientists who require access to grid resources can be created and secure access to shared medical data provided using fine grained authorization. In addition, the accountability provided by ACD makes it possible to track local users responsible for performing tasks in distributed environments in case of misuse or violation of the security policy for the VO. Indeed, the fact that ACD is based on a formal model means that it is well documented and can be certified in the future. Finally, the design of ACD is flexible enough for it to be included within the VPH Toolkit for which successful integration with AHE leads the way; its integration with IMENSE will continue to be developed in a major new project called ‘p-medicine’ (EU-FP7-270089). Support for different types of credentials such as Kerberos and Shibboleth is planned in future work which will give end users more options to choose from.

The ACD software will be available free of charge via the VPH Toolkit (toolkit.vph-noe.eu/) and will feature in future releases of the AHE that will also be distributed via the VPH Toolkit.

The authors would like to thank Prof. Dr Norbert Graf and Prof. Dr Rainer Böhle (University of Saarland) for helpful discussions on acquiring and transferring patient data to IMENSE. The authors also wish to thank Prof. Dr Nikolaus Forgó (Leibniz University, Hannover) for helpful discussions on patient data protection and data security law. We are also grateful to Nancy Wilkins-Diehr (TeraGrid), Gavin Pringle (DEISA) and David Wallom (UK NGS) for giving us permission to deploy ACD on their grid infrastructures. This work has been supported by EPSRC through the User-Friendly Authentication and Authorisation Security for Grid Environments [32] (EP/D051754/1) and RealityGrid Platform (EP/C536452/1) grants, as well as the EU FP7 ContraCancrum Project (EU-FP7-223979) [27] and Virtual Physiological Human Network of Excellence (FP7-2007-IST-223920) grants.

REFERENCES


