Data-Centric Privacy Protocol for Intensive Care Grids

Jesus Luna, Marios Dikaiakos, Manolis Marazakis, and Theodoros Kyprianou

***Abstract*—Modern e-Health systems require advanced computing and storage capabilities, leading to the adoption of technologies like the grid and giving birth to novel *health grid* systems. In particular, intensive care medicine uses this paradigm when facing a high flow of data coming from intensive care unit's (ICU) inpatients just like demonstrated by the ICGrid system prototyped by the University of Cyprus. Unfortunately, moving an ICU patient's data from the *traditionally isolated* hospital's computing facilities to data grids via public networks (i.e., the Internet) makes it imperative to establish an integral and standardized security solution to avoid common attacks on the data and metadata being managed. Particular emphasis must be put on the patient's personal data, the protection of which is required by legislations in many countries of the European Union and the world in general. In this paper, we extend our previous research with the following contributions: 1) a mandatory access control model to protect patient's metadata; 2) a major security revision to our previously proposed privacy protocol by contributing with a "quality of security" quantitative metric to improve fragmented data's assurance; and finally, 3) a set of early results to demonstrate that our protocol not only improves a patient personal data's security and privacy but also achieves a performance comparable with existing approaches.**

***Index Terms*—Data fragmentation, health grids, intensive care grid (ICGrid), privacy, security.**

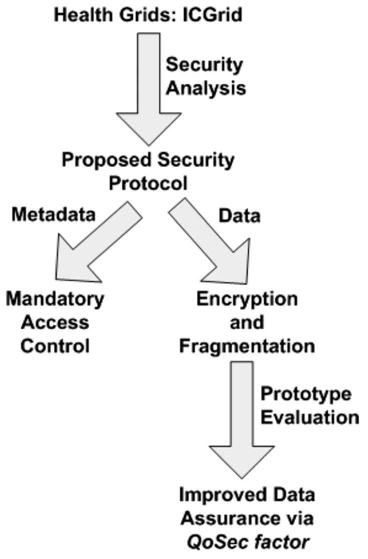

Health Grids: ICGrid
Security Analysis
Proposed Security Protocol
Metadata
Data
Mandatory Access Control
Encryption and Fragmentation
Prototype Evaluation
Improved Data Assurance via *QoSec factor*

Fig. 1. Outline of the research presented in this paper.

I. Introduction

THE data grid is becoming a new paradigm for e-Health systems due to its enormous storage potential using decentralized resources managed by different organizations. The storage capabilities in these novel "health grids" have proved quite suitable for capturing, storing and managing clinical data, and metadata [i.e., from intensive care units (ICU)], just as demonstrated by the ICGrid system prototyped by the University of Cyprus [1].

However, the health grid paradigm depends on sets of widely distributed storage elements (SEs); therefore, requiring new security mechanisms able to avoid potential leaks and the modification or destruction of stored data under the presence of external or internal attacks. In particular, patient's personal information (often referred as metadata) must be carefully guarded just as mandated by National Data Protection Legislations (i.e., in the European Union, Member States must adhere to [2]). Despite the integration of well-known privacy enhancing mechanisms in several health grid projects, previous research [3] showed the existence of security gaps at the storage level, this representing a major threat for a patient's personal information.

The work presented in this paper extends our previous research by using the methodology shown in Fig. 1: first, in order to setup our initial context, we summarize the different security gaps related with ICGrid's metadata and data just as found in our previous paper [4]. Second, in order to avoid identified data and metadata attacks (leakage, change, or destruction), while at rest into the *untrusted* SEs, we have greatly improved the previously proposed security protocol (also in [4]) in order to protect also metadata [using a mandatory access control (MAC)] and data (via fragmentation and encryption). The latter mechanisms were found to provide important security and performance guarantees, even when considering potential attackers with full control of a subset of SEs. Third and final, an important security improvement is also contributed by our research due to the analysis of the close relationship among the stored data's assurance, the selected fragmentation scheme and the grid SE's *quantitative security level*. At our best knowledge, this is a *novel security metric for health grids* and is proposed as an extension to the

Manuscript received November 14, 2009; revised June 15, 2010; accepted August 28, 2010. Date of publication; date of current version. This work was supported in part by the European Union through project CoreGRID under Contract 004265 and in part by Enabling Grids for E-sciencE-II under Contract INFSO-RI-031688.

J. Luna is with the Barcelona Digital Centre Tecnologic, 08018 Barcelona, Spain (e-mail: jluna@bdigital.org).

M. Dikaiakos is with the Department of Computer Science, University of Cyprus, CY-1678 Nicosia, Cyprus (e-mail: mdd@cs.ucy.ac.cy).

M. Marazakis is with the Institute of Computer Science, Foundation for Research and Technology—Hellas, GR-71110 Heraklion, Greece (e-mail: maraz@ics.forth.gr).

T. Kyprianou is with the Intensive Care Unit, Nicosia General Hospital, Nicosia 2042, Cyprus (e-mail: drtheo@cytanet.com.cy).

Digital Object Identifier 10.1109/TITB.2010.2073478

more generic methodology developed in [5], but that has proved useful for desktop grid systems [6].

The rest of this paper is organized as follows: basic background security technologies and legal approaches used for health grids are presented in Section II. Section III reviews the terminology related with intensive care medicine and the ICGrid system used along this paper. Then, Section IV summarizes our previous findings related with a security analysis of the former from a data-centric point of view. Section V describes in detail the contributed privacy protocol along with the middleware architecture required to implement it into ICGrid. Experimental results on the performance achieved by our proposal are shown in Section VI, while analytical results on a contributed security improvement will be presented in Section VII. Section VIII surveys and compares the state-of-the-art related with our research. Finally, Section IX presents our conclusions and future work.

II. Background on Privacy for Health Grids

As mentioned in Section I, comprehensive privacy solutions for health grids need the synergy of two different factors: legislation and technology.

A. Legal Aspects

A major concern in e-Health is adequate confidentiality of the individual records being managed electronically. The core component of many e-Health systems is the electronic health record (EHR), which is basically the patient's health record in digital format. Nowadays, EHR protection is the focus of privacy legislations around the globe. In U.S., this class of information is referred to as protected health information (PHI) and its management is addressed under the Health Insurance Portability and Accountability Act (HIPAA) [7] as well as many state laws. It has been commented that the inclusion of HIPAA authorization within the informed consent process may raise concerns about privacy [8].

In the European Union, several directives of the European Parliament and of the Council protect the processing and free movement of the EHR. The common factor of all these initiatives is the EU directive on data protection [2], which provides the general framework for the protection of privacy with respect to the processing of personal data in its widest sense. This directive goes further than the protection of the intimacy of the natural persons, since it defines *personal data* as all data related to an individual person's private, public, or professional life. However, the European Working Party on data protection, which was established under article 29 of the directive [2] and comprises all national data protection authorities of EU Member States, has recently acknowledged that some special rules may need to be adopted for key e-Health applications.

A common term referenced in current e-Health legislations is the concept of *consent*. Such consent is defined as any unambiguous, freely given, specific, and informed indication of the patient's wishes by which he agrees to the processing of his personal data. In other words, *a patient's consent enables the legal processing of his EHR*. However, what happens if, for instance, after an accident the patient is unable to give his consent for accessing his personal data at the ICU? Most of the legal issues and ambiguities related with e-Health regulations are being carefully studied; in the particular case of the European Union, the European Health Management Association (EHMA) along with the Commission established the "Legally e-Health" [9] project to study these. This study gives three basic recommendations regarding the protection of patients' data, using the terminology from RFC 2196 (see [10, Sec. IV]). These recommendations can be mapped to the *security services* shown in Table I.

TABLE I
Security Requirements for Implementing Data Protection Legislations in e-Health Environments

Legal Issue	Security Requirement	Example
Patient's Consent	Authentication, Non-repudiation, Integrity	A patient must be confidently identified (authentication) before signing an agreement (non-repudiation) allowing processing his EHR. The signed document should not be modified afterwards (integrity) without notifying the patient.
Specified Purpose	Authorization, Confidentiality, Integrity	If a patient has given his consent to re-use parts of his demographic data for statistical purposes, then a pharmaceutical company should not have access to this information (authorization) and even the personnel authorized to process these statistics should not be able to disclose i.e. the patient's name (confidentiality) or modify the record at all (integrity).

This information will be used later in this paper toward implementing a comprehensive and harmonized privacy solution for the ICGrid to be presented in Section III.

B. Technological Aspects

Enforcing privacy of patient data in health grids have spawned the development of a broad range of mechanisms. Two of these are particularly important for our research because of their wide use: the grid security infrastructure (GSI) and the electronic health card.

1) Grid Security Infrastructure: The GSI [11] is comprised of a set of protocols, libraries, and tools that allow users and applications to securely access grid resources via well-defined authentication and authorization mechanisms. In the first case, the grid client (GC) simply uses an X.509 end entity certificate to secure messages and authenticate itself to the grid service. On the other hand, for authorization purposes, GSI can use Extensible Markup Language-based protocols to retrieve security assertions from third-party services to enable features like role-based authorization. One of these third-party grid authorization services, widely used in the Enabling Grids for E-sciencE (EGEE) grid infrastructure [12], is the *virtual organization membership service* (VOMS) [13]: an attribute authority that exposes attributes and encodes the position of the holder inside the VO. Despite its functionalities, nowadays, grid authentication and authorization systems are unable to enforce access control close to the SEs and the data itself, in other words, an attacker passing over these security mechanisms (i.e., using a local account with administrative privileges or accessing physically the disks) will have full control over the stored data. These vulnerabilities are analyzed in Section IV.

2) Electronic Health Card: Smart card technology is recognized as a feasible option to enhance e-Health security, in particular authentication, confidentiality and nonrepudiation (see [14] and [15]). European Union's Member States have begun testing the electronic health card [16], a new health card that contains basic patient data, such as name, age, insurance details, and electronic prescriptions. The card includes also physical features to identify the owner, i.e., a photograph and human-readable information. With time, this card will replace EU's existing health insurance cards. Basically, this card is a smartcard that stores information in a microchip supporting authentication, authorization, and even digital signature creation. Data protection issues were critical in the design of electronic health cards; therefore, patients must be able to rely on maximum security and confidentiality, while operating smoothly in practice. A comprehensive security concept secures the protection of particularly sensitive data, so with few exceptions, the health card can only be used in conjunction with an *electronic health professional card*, which carries a "qualified" electronic signature (one that meets strict statutory criteria for electronic signatures).

Electronic health cards and smart cards, in general, represent a big step toward creating a citizen-centered health system, but despite its security advantages, internal storage space is quite limited (just few kilobytes). Thus, the use of the card must rely on external storage services over which the card cannot offer protection mechanisms. The next section will analyze in detail these security gaps.

III. ICGrid System

In this section, we introduce the required background and the respective terminology for intensive care medicine, which is the basis of the ICGrid system analyzed in this paper.

A. Intensive Care Medicine

An ICU is the only environment in clinical medicine, where all patients are monitored closely and in detail for extended periods of time, using different types of *medical monitoring devices (MMD)*. An MMD may be defined as a collection of sensors that acquire the patients' physiological parameters and transform them into comprehensible numbers, figures, waveforms, images, or sounds. Taking clinical decisions for ICU patients based on monitoring can be a very demanding and complex task requiring thorough analysis of the clinical data provided: *even the most skilled physicians are often overwhelmed by huge volumes of data, a case that may lead to errors, or may cause some form of life-threatening situation* [17]. Providing systems that actively learn from previously stored data, and suggest diagnosis and prognosis is a problem that, to our knowledge, has been overlooked in previous intensive care medicine research.

Traditionally, medical research is guided by either the concept of patients' similarities (clinical syndromes and groups of patients) or dissimilarities (genetic predisposition and case studies). Clinical practice also involves the application of commonly (globally) accepted diagnostic/therapeutic rules (*evidence-based medicine* [18]) as well as *case-tailored approaches*, which can vary from country to country, from hospital to hospital, or even from physician to physician within the same hospital. These different approaches in treating similar incidents produce knowledge, which, most of the times, remains a personal/local expertise, not documented in detail and not tested against other similar data. Global sharing of this cumulative national/international experience would be an important contribution to clinical medicine in the sense that one would be able to examine and follow up implementation of and adherence to guidelines as well as to get the benefit of sharing outstanding experience from physicians.

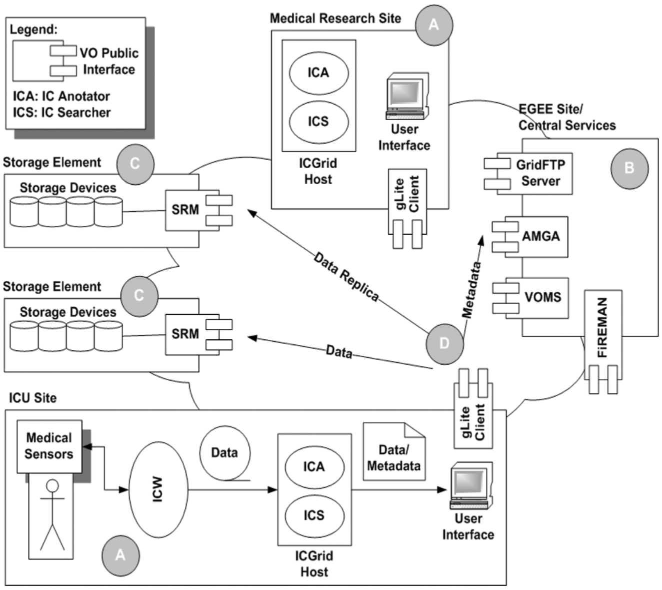


Fig. 2. Architecture of an ICGrid's VO.

B. ICGrid: Data and Metadata Architecture

Although a number of dedicated and commercially available information systems have been proposed for use in ICUs [19], which support real-time data acquisition, data validation and storage, analysis of data, and reporting and charting of the findings, none of these systems was appropriate in our application context due to the large storage requirements. Suppose that each sensor is acquiring data for storage and processing at a rate of 50 bytes per second (it is stored as text) and that there are 100 hospitals with ten beds each, where each bed has 100 sensors. Assuming that each bed is used for 2 h per day, the data collected amounts to 33.5275 GB per day. But this number only represents the data from the sensors. Additional information includes metadata, images, etc. Because grids represent a promising venue for addressing the challenges described earlier, the ICGrid system [1] has been prototyped over the EGEE infrastructure (Enabling Grids for E-sciencE [12]) and gLite [20], which is EGEE's middleware. ICGrid is based on a hybrid architecture that combines a heterogeneous set of monitors that sense the inpatients and three grid-enabled software tools that support the storage, processing, and information-sharing tasks.

The diagram of Fig. 2 represents the high-level architecture of the ICGrid system, which comprises the acquisition and annotation of parameters of an inpatient at an ICU site (bottom left) and the transfer of data replicas to two *SEs*. The transfer comprises the actual sensor data, denoted as *data*, and the information, which is provided by physicians during the annotation phase, denoted as *metadata*. We utilize the notion of a *clinically interesting episode (CIE)* to refer to the captured sensor data along with the metadata that is added by the physician to annotate all the events of interest.

When ICGrid data and metadata are transferred to SEs and metadata servers (currently a gLite metadata Catalogue (AMGA) service [21]), respectively, a set of messages are exchanged among the different entities. In particular, we should highlight that file catalog services are being provided by File Replication MAnager (FiReMAN [22]) and, authorization mechanisms rely on the X.509 credentials issued by the VOMS [13].

IV. Security Analysis of ICGrid

From the point of view of a typical health grid system, its subsystems may be attacked in several ways. Nevertheless, for the purposes of our research on data privacy, the framework proposed in [23] and extended in [3] is used to pinpoint the main concerns linked with the security of its data and metadata. In a nutshell, the use of this framework consists of determining the basic components related with the system's security (players, attacks, security primitives, granularity of protection, and user inconvenience), so that afterward they can be summarized to clearly represent its security requirements. The security analysis framework obtained from a previous research [3] is summarized in this section in the context of the ICGrid system, considering also the underlying security mechanisms presented in Section II.

A. Identifying the Elements for the Security Analysis

As mentioned at the beginning of this section, the first step in our analysis is to identify the elements that play a security-related role in ICGrid.

1) *Players:* Referring to Fig. 2, four data readers/writers are involved: 1) the ICU and medical research sites (marked with A in the figure) that produce and consume the data; 2) the EGEE central services (marked B) that perform VO authentication and authorization, as mentioned in Section II-B; 3) the EGEE *storage facilities* for data and metadata (marked as C in figure); and finally, 4) the "wire" or WAN links (public and private) conveying information between the other players (marked as D).
2) *Attacks:* The generic attacks that may be executed over ICGrid are related with 1) adversaries on the wire; 2) revoked users using valid credentials on the central services during a period of time, while the revocation data is propagated through the grid; and 3) adversaries with *full control* of the EGEE storage facilities. Each one of these attacks may result in data being leaked, changed, or even destroyed.
3) *User inconvenience:* It is critical for IGGrid operation to have minimum latencies when reading and retrieving the stored data and metadata from the EGEE site. Since smart cards—like the electronic health card explained in Section II-B2—are beginning to be introduced into National Health Systems, it is feasible to consider that involved entities (i.e., patients and physicians) will require them for performing operations into our health grid scenario.
4) *Security primitives:* Two security operations take place within the ICGrid: a) *authentication and authorization* via GSI-like mechanisms (see Section II-B1); and b) *consent* just as explained in Section II-A.
5) *Trust assumptions:* We assume that a) the security tokens used for authentication and consent (i.e., electronic health cards) are personal, intransferable, and tamper-resistant; b) EGEE sites and/or ICU premises have full control over the data and metadata stored on them; c) data are encrypted on the public link due to secure functionalities (i.e., via SSL); and d) the EGEE central services are *trusted* because they are managed in a secure manner; therefore, providing high assurance to its operations.

TABLE II
Summary of Security Issues Related With ICGrid

	Adversary on the wire			*Revoked user w/Central Service*			*Adversary w/Storage Site*		
Damage	L	C	D	L	C	D	L	C	D
ICGrid	N	N	Y	Y	Y	Y	Y	Y	Y

B. Security Analysis Results

Based on the elements identified in the previous section and our previous work [3], Table II summarizes the vulnerabilities identified in the ICGrid system. Results are categorized by possible attacks (main columns) and types of damage—the leak (L), change (C), and destroy (D) subcolumns. Cells marked with a "Y" mean that the system (row) is vulnerable to the type of damage caused by this particular attack. Cells marked with a "N" mean that the attacks are not feasible, or cannot cause a critical damage.

From Table II, we conclude that current health grid authentication and authorization systems like the ones presented in Section II-B are unable to enforce access control close to the SEs and the data itself. In other words, an attacker that bypasses these security mechanisms (by using a local account with administrative privileges or by physical access to the disks) will have full control over the stored data. Unfortunately, merely using cryptography at the SEs is not enough because encryption keys can be leaked by a local attacker. Moreover, this approach would impose a significant performance penalty. In the following section, we introduce a data-centered protocol designed to address these particular privacy concerns.

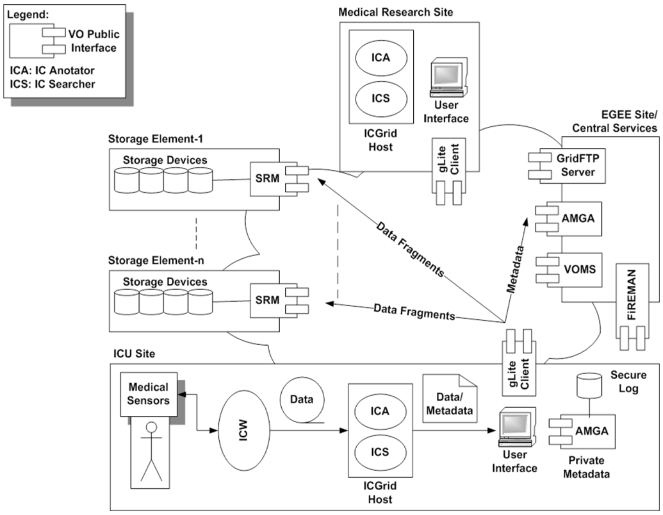


Fig. 3. Secure ICGrid architecture.

V. Introducing a Secure ICGrid: Protecting Metadata and Data

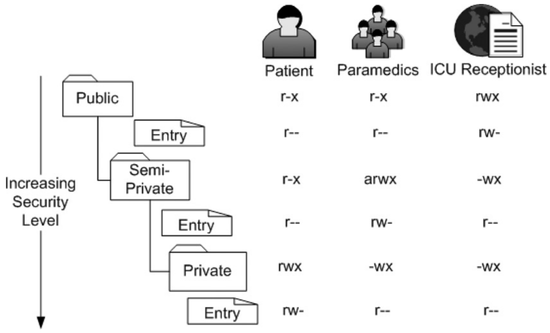
In this section, we further improve the protocol introduced in [3] by presenting the main components of an architecture proposed to provide security to the ICGrid system introduced in Section III. *The specific goal of our proposal is to avoid data and metadata attacks (leakage, change, or destruction), while at rest into the untrusted SEs.* It is worth noticing that performance issues related with the mechanisms being used have been carefully considered in our design (more about this in Section VI).

A. Architecture and Design Principles

Because our previous security analysis [3] found that health grid's metadata and data require different security policies, the enforcement mechanisms presented in the rest of this section implement a differentiated approach for metadata (see Section V-B) and data (see Section V-C) based on ICGrid's current architecture (see Fig. 2). Our improved protocol's proposal uses three basic mechanisms.

1) An information dispersal algorithm (IDA) providing *high availability and assurance* for the ICGrid by means of *data fragmentation*. In a fragmentation scheme [24], a file f is split into n fragments, all of these are signed and distributed to n SEs, one fragment per SE. The user then can reconstruct f by accessing m fragments ($m \leq n$) arbitrarily chosen.
2) A symmetric cryptosystem implemented at the SEs [i.e., via a hardware security module (HSM)], which is able to provide confidentiality to the stored data, while keeping a good balance between security and performance.
3) An MAC mechanism based on the Bell–LaPadula model [25] to protect the private metadata stored at the ICU's premises.

The overall architecture of this "Secure ICGrid" is shown in Fig. 3



	Patient	Paramedics	ICU Receptionist
Public	r-x	r-x	rwx
Entry	r--	r--	rw-
Semi-Private	r-x	arwx	-wx
Entry	r--	rw-	r--
Private	rwx	-wx	-wx
Entry	rw-	r--	r--

Fig. 4. MAC model for ICGrid's metadata.

B. Metadata Security

AMGA stores metadata in a hierarchical structure that resembles a Unix file system, also with a native authorization model based on access control lists (ACLs) [26] with POSIX-like permissions per-entry and directory ($r =$ read, $w =$ write, and $x =$ change into directory) and an additional "admin flag" allowing users in a group to administer the ACLs of an entry. Using the latter mechanism, we have defined an authorization model for ICGrid's metadata based on the following MAC rules.

1) The *simple security property* states that a subject at a given security level may not read an object at a higher security level (no read-up).
2) The **-property* (read star-property) states that a subject at a given security level must not write to any object at a lower security level (no write-down) and may only append new data to any object at a higher security level.

Bell–Lapadula's model applied to ICGrid's private metadata (implemented over an AMGA server located at the ICU's premises, just as shown also in Figs. 5 and 6) can be seen in Fig. 4. The proposed MAC model is able to provide a basic level of confidentiality to the patient's private metadata, while at the same time "protecting" him from accidentally disclosing this information to the lower security levels. In this example, we have defined three different players (patient—owner, paramedics—group, and the ICU receptionist—others) and also, three levels of authorization (public, semiprivate, and private). With the proposed AMGA's permissions on directories and entries, it is possible to achieve the following MAC.

1) *Public metadata:* Both patient and paramedics can read the entries, but only the ICU receptionist can read and write them (i.e., schedule a new visit by the physician).
2) *Semiprivate metadata:* The Paramedics can read and write entries (i.e., emergency information), the ICU receptionist can only append new ones (the paramedics group requires the admin flag to set read-only permissions to these newly created entries) and the patient is only able to read this metadata.
3) *Private metadata:* This is the most confidential level of the metadata; therefore, only the patient has full control over it (administrative permissions are implicit, since he is the owner of his directories), while paramedics and ICU receptionists only can append new entries (the patient must manage permissions of these newly created entries).

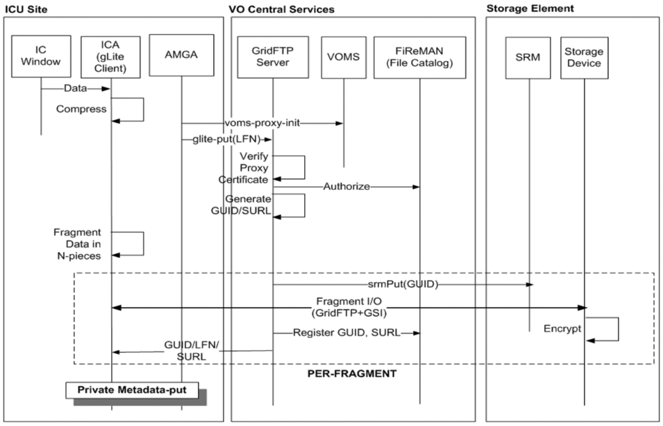


Fig. 5. Secure ICGrid: transferring data.

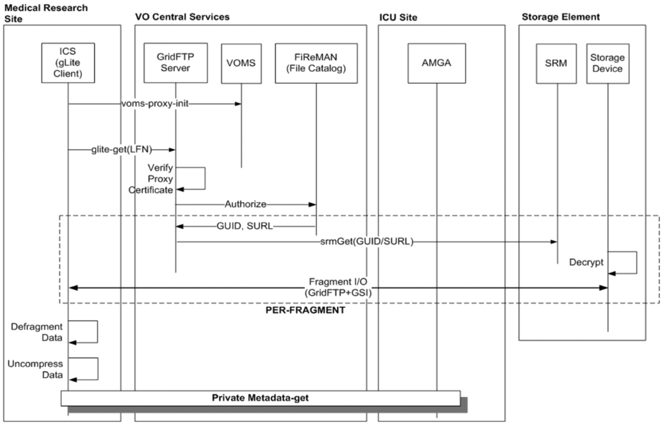


Fig. 6. Secure ICGrid: retrieving data.

Enforcing the *-property's append-only mode conveys an administrative overhead for both, patients and paramedics, which must manage permissions for entries being created by lower security subjects. Also it is worth to notice that native AMGA's authorization mechanism cannot prevent a malicious system administrator from accessing the metadata of all the stored patients; however, our belief is that the use of a final component, the *secure log* will allow auditing AMGA operations.

Our future work considers the use of cryptographic techniques to provide greater confidentiality and even a consent-like mechanism (based on electronic signatures) to AMGA's metadata. This research is introduced in Section IX.

C. Data Security

Fig. 5 shows how the different components introduced in Section V-A interact with the central services when an IC annotator (ICA) stores data into the ICGrid system. In Fig. 5, we use the file naming notation from [27], when referring to the data being managed by the grid: 1) logical file name (LFN) (a human readable identifier for a file); 2) global unique identifer (GUID) (a logical identifier, which guarantees its uniqueness by construction); and 3) site URL (SURL) (specifies a physical instance of a file replica, which is accepted by the SE's Storage Resource Manager (SRM) interface).

The core of our proposal is: 1) the IDA mechanism implemented at the ICU's gLite client responsible for fragmenting the data to be uploaded via a GridFTP client and 2) the symmetric encryption of the fragments taking place at the SEs. Notice that the file catalog must keep the exact location (SURL) of each fragment corresponding to a particular data file (GUID).

It is worth noticing that attackers with full control of a SE will be unable to compromise the entire data file unless they can retrieve and decrypt at least other $m - 1$ fragments from participating SEs. Our research also shows that this can be even more difficult to achieve if each SE is associated with a numeric value representing its actual "security level," just as presented in Section VII.

A second scenario (see Fig. 6) considers an IC searcher (ICS) retrieving data from the ICGrid: in this case, the fragments are decrypted, then transferred from *m-SEs*, and conveyed through a secure channel to the ICS GridFTP client, where the defragmentation process will take place. Note that with the proposed fragmentation approach, if at most *n–m* SEs are offline, then it will still be possible to rebuild to original data.

In the next section, we show some preliminary performance results obtained with a prototype of the protocol just shown, as a way to identify the tradeoffs to be considered in a real deployment.

VI. Experimental Results

We have setup the following test bed to measure the expected performance to be achieved with the fragmentation and encryption mechanisms of the protocol proposed in Section V-C:

1) *Grid client:* This CentOS4-based node has been configured as a "gLite user interface." It is an IBM xSeries 335 with two Intel Xeon HT processors @ 2.8 GHz and 2 GB of RAM.
2) *Storage element:* We have used for the tests a "DPM_ mysql Storage Element" running over Scientific Linux version 3.09. The SE uses a Dell PowerEdge1400, with two Intel Pentium III processors @ 800 MHz and 784 MB of RAM.

For the data we have prepared, one synthetic sample corresponding to one day of ICGrid's operation for a hospital (approximate 351 563 kB). The *gzip* utility was used with its default parameters for compression at the GC, while for SE-encryption and decryption, we used the *aes-128-cbc* algorithm[1] from the OpenSSL library (version 0.9.8g). To perform the tests under the same conditions offered by the ICU's satellite link to Internet (uplink = 125 kB/s and downlink = 125 kB/s), we installed a software-based bandwidth shaper at the GC. Finally, for the fragmentation algorithm, we implement the Reed–Solomon IDA based on open-source code from Onion Networks [28].

For comparison purposes, we measure the protocol performance as the user time (reported by the *time* utility) consumed by each phase of the following scenarios.

[1] Advanced Encryption Algorithm in Cipher Block Chaining mode, with symmetric keys of 128 bits length.

TABLE III
REPORTED SIZES (IN kB) FOR THE SYNTHETIC DATA SAMPLE AFTER COMPRESSION AND ENCRYPTION

Original	Compressed	Encrypted
351563	212125	287258

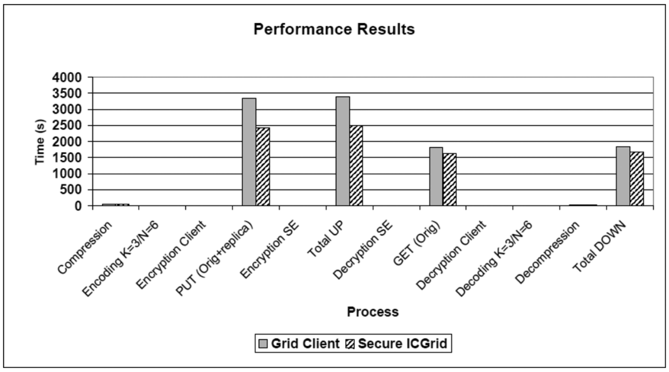


Fig. 7. Comparing the performance achieved when processing a day of ICGrid's data with the traditional approach (GC graph) and the proposed protocol (secure ICGrid graph).

1) *GC encryption:* This approach performs encryption/decryption at the GC and is commonly used by existing solutions (see Section VIII). The steps taking place are: data compression, encryption, and transfer to two SEs (to also simulate the placement of a replica) via clear-text FTP. The inverse sequence is used to retrieve it from the SE. *This scenario will be taken as the baseline for the results shown in this section.*
2) *Secure ICGrid:* This scenario simulates the basic steps proposed by our protocol, namely data compression, fragmentation (in a $m = 3$, $n = 6$ mode, which is space equivalent to replication), transfer via a GSIFTP encrypted channel to the SEs, and finally, encryption of each fragment at each one of these. The inverse sequence of steps is used to retrieve three stored fragments from the SEs, so the original file can be rebuilt.

Each scenario was executed 50 times to isolate potential overhead being caused by other processes concurrently running at the server and when possible, the experiments took advantage of parallelizing some phases of the whole process (i.e., sending and retrieving the data fragments).

Table III shows how the size of the synthetic data sample changed after the compression and encryption processes. It is worth noticing that the compressed data's size is about 60% of the original one; however, after encryption the size increased approximately by 35%.

Fig. 7 and Table IV show the average performance results and their breakdown, respectively, for the 50 planned executions of both scenarios and using the synthetic data sample. This side-by-side comparison of the GC process (*baseline*) versus the secure ICGrid process (*proposal*), allowed us to determine, which specific phases of the contrinuted protocol, in fact, convey a gain in performance and security. Aggregated values for the tested scenarios are given by the *TOTAL UP* and *TOTAL DOWN* bars.

TABLE IV
BREAKDOWN OF OBTAINED OVERHEADS (EXPRESSED IN SECONDS, N/R = NOT REQUIRED FOR THIS SCENARIO)

Stage	*Grid Client (Baseline)*	*Proposed Approach*
Compression	37.63	37.63
Encoding	n/r	11.01
Client Encryption	5.14	n/r
PUT (w/replica)	3336.77	2432
SE Encryption	n/r	4.92
SE Decryption	n/r	10.47
GET (w/o replica)	1814.43	1621
Client Decryption	8.56	n/r
Decoding	n/r	4.79
Decompression	22.85	22.85

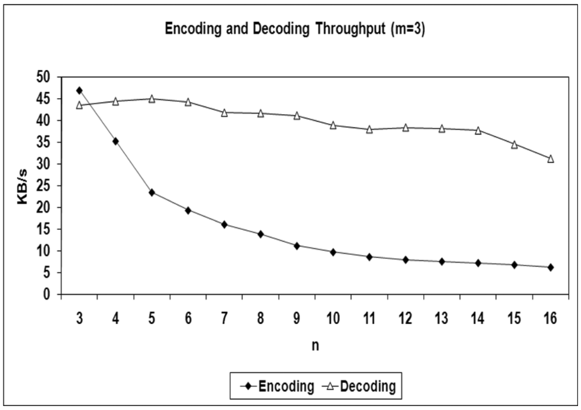


Fig. 8. Data throughput for the IDA mechanism with different values of n.

Just as expected, most of the time spent by both processes was due to the uploading and downloading of the data to and from the grid, respectively. However, in both cases, the proposed protocol performs better than the traditional approach. This behavior is a result not only of the uploaded-data size (as shown in Table III), but also of the improved bandwidth use when uploading and downloading are in parallel of the generated fragments (for the experiments, we ran three simultaneous GridFTP sessions a couple of times).

When comparing the performance of the cryptographic and the fragmentation mechanisms (see Table IV), we can observe that despite encryption costs in CPU almost half the time compared to fragmentation (approximate 5 s of the former versus 11 s of the latter), both cost much less than the uploading process. A similar observation holds when comparing the decryption and defragmentation processes, even though in this case, the latter spent almost half the time (approximate 4.8 s) of the former (approximate 8.6 s).

Finally, we run in a different subset of experiments several operation modes of the IDA to find the performance tradeoffs incurred when varying the total number of fragments (n-parameter). Figs. 8 and 9 show our results.

Although a bigger n implies a more fault-tolerant system, it is easy to observe that the encoding process involved a bigger computational cost (i.e., less data throughput). Therefore, the right choice of the m and n parameters for the IDA is quite important for the overall ICGrid operation. The "replication-like" scheme used in our experiments ($m = 3$ and $n = 6$) was a first approach for a balanced performance-security solution;

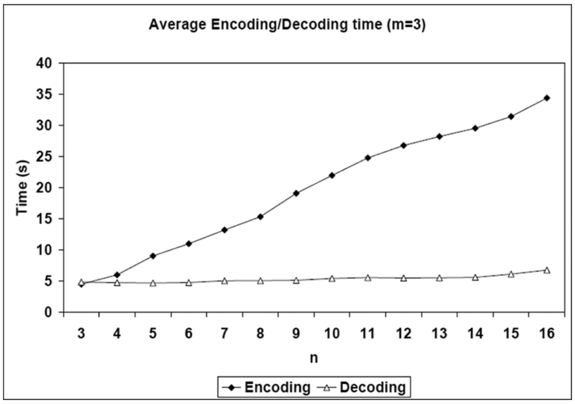


Fig. 9. Encoding and decoding times for the IDA mechanism with different values of n.

however, we have found that another criteria for selecting these values may take into account the close relationship among the stored data assurance, the IDA parameters, and the SE's security level. The next section will prove this rationale in more detail.

VII. Improved Assurance for Stored Data

While developing the research presented in previous sections, we realized that SEs available on the grid infrastructure were heterogeneous, both in hardware and software, although in most of the cases, there were some minimum requirements to be fulfilled before allowing them to interoperate. Therefore, a question arises: if two SEs are heterogeneous, to what extent is the data assurance they provide also different? This was an important question to solve before deploying the proposed security protocol for ICGrid.

Our hypothesis was that if subsets of SEs with analogous features were clearly identified and their security level quantified, then it could be possible not only to choose the adequate IDA to use but also an ICGrid user should be able to request storing its data *only* at those sites fulfilling some minimum guarantees or *quality of security (QoSec)* level. To achieve this goal, we propose a policy-based approach, where 1) each SE is associated with a policy modeling its security behavior and, 2) an evaluation methodology for computing a *quantitative* QoSec metric associated with such policy.

In prior research [29], we used the certificate policy [30] associated with a grid user's X.509 certificate to perform a quantitative security evaluation. We believe an analogous approach for SEs is not only feasible but also that because this evaluation is based on quasi-static parameters (i.e., an established security policy), it should not add considerable overhead to the protocol proposed in Section V.

Once evaluated, a numeric QoSec factor can be associated with every SE from ICGrid's VO; moreover, a proprietary set of reference levels can be established to group these discrete values into a major set, for example, (L_0, L_1, L_2, L_3), which represents four different QoSec. Any ICGrid client uploading data will request a minimum QoSec to be fulfilled by the SEs.

We find two main challenges with the proposed approach.

1) The QoSec evaluation.
2) The definition and auditing of the SE's security policy.

The rest of this section will present our work to cope with both challenges, using the reference evaluation methodology (REM) for the former and, establishing a "controlled" scenario based on a *Grid Policy Management Authority* (PMA) for the latter. We present analytical results that show the existing relationship among data assurance, QoSec and the data fragmentation scheme being used.

A. REM in a Glimpse

The core of the proposal to quantify the QoSec associated with a SE is the REM introduced in [5], which has been successfully applied in other security-related environments, i.e., public key infrastructures (PKI). Basically, the REM takes a policy to evaluate, formalize it to ease the further evaluation step over an homogeneous metric space, uses a set of reference levels to apply a distance criteria, and finally, obtains a number that corresponds to the policy's security level. The rest of this section presents an illustrative example that shows basic use of the REM methodology to evaluate a single security provision for an SE. Section VII-B shows the evaluation of a full-security policy (a set of individual provisions). Interested readers that would like to take a closer look at the details and formalisms behind REM should refer to [31].

Step 1 (Policy formalization): In REM, a security policy is formed via a set of individual security provisions, so let us suppose that a SE's security policy contains the provision called $P_x =$ use of cryptographic mechanisms." To formalize this provision, it is necessary to state the possible values that are allowed for it to take into the VO. For our example, let us suppose the following values for P_x (from least to most secure) "none, software-based, network-based, and disk controller-based," this means that P_x has a cardinality of 4.

Step 2 (Evaluation technique): If a particular SE uses the cryptographic mechanisms provided by a network-based HSM, then the provision P_x takes as value the *n-tuple* $(1, 1, 1, 0)$, which means that it is more secure than a software mechanism [represented by $(1, 1, 0, 0)$], but less than one implemented on the disk controller $(1, 1, 1, 1)$. Using REM's nomenclature, this implies a *local security level* or $\text{LSL}_x = 3$. It is easy to see that after evaluating a whole formalized security policy for a particular SE, we will have a set of these tuples or in other words, a *matrix* M_x.

To obtain the *global security level* (*QoSec*) for this particular SE, it is necessary to have a new zero-matrix called M_ϕ, which contains only 0's as its elements. Notice that M_ϕ represents the least secure SE that can be found into an ICGrid VO. The QoSec in our example is defined as the Euclidean distance between M_x and M_ϕ, i.e.,

$$d(M_x, M_\phi) = \sqrt{\sigma(M_x - M_\phi, M_x - M_\phi)} \tag{1}$$

where

$$\sigma(M_x - M_\phi, M_x - M_\phi) = Tr\left((M_x - M_\phi)(M_x - M_\phi)^T\right). \tag{2}$$

Obviously, the REM technique considers more complex scenarios, where, for example, different provisions have different security weights (some provisions are more important than others) and even different cardinalities [31].

B. QoSec Evaluation: Analytical Results

We used the model presented in [32] to evaluate the *assurance* of the data stored in a SE with the proposed security protocol. *Data assurance* is defined as the probability that the ICGrid data will not be compromised under the assumption that its SEs were target of a successful attack. Formally speaking, we assume that an ICGrid VO with N-SEs is built in such a way that the same configuration (and therefore, the same *security level*) is present in at most $\lceil \lambda N \rceil$ of them, where $\lambda \in \mathcal{R}$ and $0 < \lambda < 1$. A potential attacker could then compromise at most $\lceil \lambda N \rceil$ SEs, because the same vulnerabilities are to be found in all of them. Therefore, it makes sense to state that an ICGrid VO can be considered *secure* if it is composed of heterogeneous SEs, where of course $\lambda \approx 0$.

According to Mei *et al.* [32], this latter fact can be represented by the conditional probability of the event "at most m–1 SEs, each storing an encrypted fragment of f, are down or compromised" given the event "ICGrid has been attacked." This probability, known as the *distribution assurance* $A_{\phi}(\mu)$ for a dispersal algorithm μ applied over a data file f, is denoted as following:

$$A_{\phi}(\mu) = 1 - \sum_{i=m}^{n} \frac{\binom{\lceil \lambda N \rceil}{i} \binom{\lfloor N - \lambda N \rfloor}{n-i}}{\binom{N}{n}}. \tag{3}$$

According to the analytical model (3), it is more likely to find *subsets* of $\lceil \lambda N \rceil$-SEs sharing the same configuration. Thus, higher data assurance is obtained with more secure (QoSec $\rightarrow \infty$) *or* heterogeneous ($\lambda \rightarrow 0$) SEs. Therefore,

$$\text{QoSec} \propto \frac{1}{\lambda} \Rightarrow \text{QoSec} = \frac{k}{\lambda}. \tag{4}$$

We can rewrite (3) in terms of the proposed *QoSec* with the factor $k = 1$:

$$A_{\phi}(\mu) = 1 - \sum_{i=m}^{n} \frac{\binom{\lceil \frac{N}{\text{QoSec}} \rceil}{i} \binom{\lfloor N - \frac{N}{\text{QoSec}} \rfloor}{n-i}}{\binom{N}{n}}. \tag{5}$$

Next, we examine an hypothetical ICGrid VO spanning over SEs from three different institutions: Greece's HellasGrid, CERN's LCG, and Spain's IRISGrid. Currently, each of these installations is a member of the same Grid PMA, which means that they fulfill a minimum set of security rules (provisions) established—and audited—by the European Grid PMA (EUGridPMA) according to its "authentication profile" document [33]. As a first step, we applied the REM technique to compute the numeric QoSec associated with the certificate policy [34] subsection called "technical security controls"—mostly applicable to the SEs of this hypothetical VO—of HellasGrid (QoSec = 4.47), CERN (QoSec = 6.00), IRISGrid (QoSec = 5.48), and the minimum reference EUGridPMA (QoSec = 4.24).

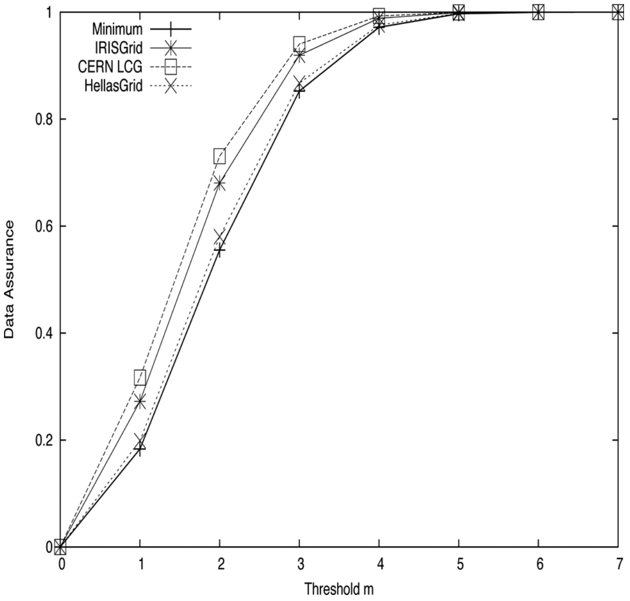


Fig. 10. Relationship between data assurance and QoSec for ICGrid's security protocol. The horizontal axis represents the number of fragments (m) to be retrieved from the SEs for rebuilding the original data.

The relationship between the assurance of the encrypted data at rest and the computed QoSec can be obtained by plotting expression (5), just as seen in Fig. 10. For the evaluation, we considered an ICGrid VO with 100 SEs ($N = 100$), implementing the data fragmentation scheme from Section VI, where ($n = 6$).

The figure shows that, despite the small difference between the computed QoSecs, CERN's nodes achieve the greatest data assurance (≈1) with the smallest number of fragments. On the other hand, a grid installation fulfilling only the minimum security requirements (QoSec = 4.24) requires the retrieval of more fragments from the SEs in order to fully defragment the stored data with the same assurance. Also notice that for the fragmentation scheme used in Section VI ($n = 6$ and $m = 3$), the best data assurance is obtained by the SE with the highest QoSec.

The results presented in this section are a first approach to compute the SE's QoSec using a subset of provisions from a certification policy; however, in a real-world installation, we should evaluate a comprehensive security policy specifically built for this resource. Section IX will return to this topic, while outlining our on-going research.

VIII. Related Work

Nowadays, most of the work related with health grids' security and privacy focuses on "high-level" authentication and authorization mechanisms that rely on Grid-IDs and VOMS-like infrastructures [13]; therefore, *leaving data vulnerable in the untrusted SEs*. An example of these kind of mechanisms can be seen, i.e., in the Health-e-Child [35], Mammogrid [36], Neugrid [37], and MEDICUS [38] health grids. Also the technological roadmaps BRIDGES [39] and SHARE [40] consider similar mechanisms.

The research that is closely related with the work presented in this paper has been presented in [41], where Montagnat *et al.* also used the gLite middleware to protect medical images. Their system ensures medical data protection through data access control, anonymization, and encryption. A fundamental difference with our approach is the use of encryption at the GC, which requires retrieving the encryption key from a hydra keystore for decrypting the image. With our research, it has been shown that such approach does not only introduce uncertainties about the key's confidentiality (it may be compromised at the GC), but also has a performance lower than our proposal (as seen in Section VI). About fragmentation, Montagnat *et al.* [41] use it only at the keystore, but does not consider its advantages for the data itself (high availability).

There are other state-of-the-art distributed storage systems that, even though they have not been specifically designed for the health grid, they have focused on low-level data protection by solely implementing cryptographic mechanisms. For example, in OceanStore [42], stored data are protected with redundancy and cryptographic mechanisms. An interesting feature in OceanStore is the ability to perform server-side operations directly on the encrypted data, this increases system's performance without sacrificing security. However, it is worth to mention the Farsite system [43], which provides security and high availability by storing encrypted replicas of each file on multiple machines. Just as mentioned along this paper, sole use of cryptography might not offer enough protection against an attacker with full control of the storage device (therefore, being able to obtain the encryption key, i.e., via a memory dump).

A second group of systems relying on data fragmentation include POTSHARDS [44], which implements an storage system for long-time archiving that does not use encryption, but a mechanism called "probably secure secret splitting" that fragments the file to store prior to distributing it across separately managed archives. A similar approach is given by Cleversafe [45] via an IDA (based on the Reed–Solomon algorithm) for its open-source *dispersed storage project*. In general, both POTSHARDS and Cleversafe are interesting solutions that solve the management problems posed by cryptosystems and long-living data; however, we have shown that security levels achieved only by fragmenting the files may not be strong enough for some highly sensitive environments.

From surveyed state-of-the-art, we could not find any reference to works considering the use of mechanisms analogous to the proposed *QoSec* for improving the assurance of stored data.

IX. Conclusion and Future Works

In this paper, we have presented the final part of our research on data-level security for health grids. After summarizing previous work about security requirements of the proposed intensive care medicine scenario (the ICGrid system), we concluded the need to provide different levels of protection for metadata and data in order to mitigate vulnerabilities found with untrusted SEs and GC that could compromise sensitive material (i.e., patient's personal data).

The second part of this paper extended our previous research by contributing a privacy protocol aimed to protect the metadata and data, using a MAC for the former and cryptography with fragmentation for the latter. Despite its simplicity, the proposed metadata approach was able to enforce different levels of authorization for a patient's personal data, in compliance with the e-Health Legislations previously surveyed. On the other hand, data protection was found to require *a mechanism* able to offer security guarantees despite attackers taking full control of a subset of SE. Such mechanism was, in fact, comprised of two approaches: symmetric encryption and data fragmentation. Even though the use of cryptography and fragmentation is not new for securing data in distributed systems, a contribution of our research in this field consisted of enhancing these mechanisms with the adoption of a numeric *QoSec* level representing the guarantees offered by the SEs to the grid user's stored data. At our best knowledge, this is a novel contribution for the health grid topic. Applying an analytical model, we have shown for three production grids: the strong relationship between the data assurance, the number of fragments required to retrieve a file, and the SE's QoSec. Therefore, we strongly believe in the security advantages of using the QoSec factor as an input to the IDA.

As a proof of concept, the architecture required to support the proposed protocol was prototyped using components entirely from the gLite middleware, for example, the proposed MAC was modeled via AMGA's ACLs. The use of cryptography with fragmentation for ICGrid's data was justified—from a performance point of view—with a SE able to encrypt the fragments coming from a prototyped client. Our experiments found that data transfer operations (upload and download) contribute with most of the protocol's overhead; therefore, suggesting us to keep transferred data as small as possible. Taking into account that the encrypted data is greater in size than its clear-text counterpart, we highly recommend not performing encryption at the GC, contrary to existing state-of-the-art implementations. Notice that the use of fragmentation greatly improves overall security because potential attackers, even with full control of one SE, would need to compromise more SEs to compromise an ICGrid file.

Our future work goes in the direction of defining a policy for SEs, beyond the certification policy used in this paper, able to model with confidence their security properties, and the grid user's expectation from the authentication, authorization, confidentiality, integrity, privacy, and availability point of view. We are also planning to study, along with AMGA's creators, the repercussions of using encryption at different levels of the metadata. For ICGrid's authentication purposes, we will begin researching smartcard-based approaches (i.e., [15]); despite their storage constrains, these devices might be useful for securely keeping encryption keys.

Acknowledgment

The authors would like to thank Prof. A. Bilas (Foundation for Research and Technology—Hellas) for his valuable review and Mr. A. Katsifodimos (University of Cyprus) for his technological support.

References

[1] K. Gjermundrod, M. Dikaiakos, D. Zeinalipour-Yazti, G. Panayi, and T. Kyprianou, "Icgrid: Enabling intensive care medical research on the egee grid," in *From Genes to Personalized HealthCare: Grid Solutons for the Life Sciences. Proceedings of HealthGrid 2007*, Amsterdam, The Netherlands, IOS Press, 2010, pp. 248–257.

[2] Various, "EU Parliament. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data," Official Journal of the European Communities of 23 November 1995 No L. 281 p. 31., Oct. 1995, 2010.

[3] J. Luna, M. Flouris, M. Marazakis, and A. Bilas, "An analysis of security services in grid storage systems," in *Proc. CoreGRID Workshop Grid Middleware 2007*, Jun., pp. 171–185.

[4] J. Luna, M. Flouris, M. Marazakis, A. Bilas, M. Dikaiakos, H. Gjermundrod, and T. Kyprianou, "A data-centric security analysis of icgrid," in *Proc. CoreGRID Integr. Res. Grid Comput.*, 2008, pp. 165–176.

[5] V. Casola, R. Preziosi, M. Rak, and L. Troiano, "A reference model for security level evaluation: Policy and fuzzy techniques," *J. UCS*, vol. 11, no. 1, pp. 150–174, 2005.

[6] J. Luna, M. Flouris, M. Marazakis, and A. Bilas, "Providing security to the desktop data grid," in *Proc. IEEE IPDPS 2008*, Miami, FL, Apr. 14–18, pp. 1–8.

[7] Online. (1996). "U.S. Department of Health and Human Services: HIPAA Law," [Online]. Available: http://www.cms.hhs.gov/HIPAAGenInfo/Downloads/HIPAALaw.pdf

[8] A. L. Dunlop, T. Graham, Z. Leroy, K. Glanz, and B. Dunlop, "The impact of hipaa authorization on willingness to participate in clinical research," in *Proc. Ann. Epidemiol.*, Nov. 2007, vol. 17, pp. 899–905.

[9] Various. (2006, Jan.). "European Health Management Association Legally eHealth-Deliverable 2," [Online]. Available: http://www.ehma.org/_fileupload/Downloads/Legally_eHealth-Del_02-data_Protection-v08(revised_after_submission).pdf, processing Medical data: data protection, confidentiallity and security.

[10] B. Fraser, "Site Security Handbook," RFC 2196 (Informational), Sep. 1997, [Online]. Available: http://www.ietf.org/rfc/rfc2196.txt

[11] V. Welch. (2005). "Globus toolkit version 4 grid security infrastructure: A standards perspective," [Online]. Available: http://www.globus.org/toolkit/docs/4.0/security/GT4-GSI-Overview.pdf, the Globus Security Team.

[12] Online. (2008). "Enabling Grids for E-SciencE project," [Online]. Available: http://www.eu-egee.org/

[13] R. Alfieri, R. Cecchini, V. Ciaschini, L. dellAgnello, A. Frohner, A. Gianoli, K. Lorentey, and F. Spataro, "VOMS, an authorization system for virtual organizations," in *Proc. First Eur. Across Grids Conf.*, Feb. 2003, pp. 13–14.

[14] D. Jones, "Smart cards—The key to secure and flexible healthcare provision," *Card Technol. Today*, vol. 15, no. 11, p. 8, 2003.

[15] W. B. Lee and C. D. Lee, "A cryptographic key management solution for hipaa privacy/security regulations," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 1, pp. 34–41, Jan. 2008.

[16] Various. (2006, Oct.). "Federal ministry of health: The electronic health card," [Online]. Available: http://www.die-gesundheitskarte.de/, public Relations Section. Berlin, Germany.

[17] B. Hayes, "Guardian: A prototype intelligent agent for intensive care monitoring," *Artif. Intell. Med.*, vol. 4, pp. 165–185, 1992.

[18] D. Sackett, *Evidence-Based Medicine: How to Practice and Teach EBM*, 2nd ed. Maryland Heights, MO: Churchill Livingstone, 2000.

[19] B. Dawant, "Knowledge-based systems for intelligent patient monitoring management in critical care environments," in *Biomedical Engineering Handbook*, J. D. Bronzino, Ed. Boca Raton, FL: CRC Press, 2000.

[20] Online. (2008). "glite: Lightweight middleware for grid computing," [Online]. Available: http://www.glite.org/

[21] N. Santos and B. Koblitz, "Distributed metadata with the AMGA metadata Catalog," in *Proc. Workshop Next-Gener. Distrib. Data Manage. HPDC-15*, Jun. 2006, pp. 1–6.

[22] Online. (2005). "EGEE: FiReMAN Catalog User Guide," [Online]. Available: https://edms.cern.ch/document/570780

[23] E. Riedel, M. Kallahalla, and R. Swaminathan, "A framework for evaluating storage system security," in *FAST*, D. D. E. Long, Ed. Berkeley, CA: USENIX, 2002, pp. 15–30.

[24] M. O. Rabin, "Efficient dispersal of information for security, load balancing, and fault tolerance," *J. ACM*, vol. 36, no. 2, pp. 335–348, 1989.

[25] D. E. Bell and L. J. LaPadula, "Secure computer systems: A mathematical model, volume ii," *J. Comput. Security*, vol. 4, no. 2/3, pp. 229–263, 1996.

[26] Online. (2008). "Amga: Users, groups and acls," [Online]. Available: http://project-arda-dev.web.cern.ch/project-arda-dev/metadata/groups_and_acls.html

[27] S. Burke, S. Campana, A. D. Peris, F. Donno, P. M. Lorenzo, R. Santinelli, and A. Sciaba. (2006, Nov.). "gLite 3.0 User Guide," [Online]. Available: http://glite.web.cern.ch/glite/documentation/

[28] Various. (accessed 2008, Jun.). "Onion Networks," [Online]. Available: http://www.onionnetworks.com/developers/

[29] V. Casola, N. Mazzocca, J. Luna, O. Manso, and M. Medina, "Static evaluation of certificate policies for grid pkis interoperability," in *Proc. 2nd Int. Conf. ARES 2007*. Washington, DC: IEEE Computer Society, pp. 391–399.

[30] S. Chokhani, W. Ford, R. Sabett, C. Merrill, and S. Wu. (2003, Nov.). "Internet X.509 public key infrastructure—Certificate policy and certification practices framework," RFC 3647 (Informational), [Online]. Available: http://www.ietf.org/rfc/rfc3647.txt

[31] V. Casola, A. Mazzeo, N. Mazzocca, and V. Vittorini, "A policy-based methodology for security evaluation: A security metric for public key infrastructures," *J. Comput. Security*, vol. 15, no. 2, pp. 197–229, 2007.

[32] A. Mei, L. V. Mancini, and S. Jajodia, "Secure dynamic fragment and replica allocation in large-scale distributed file systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 14, no. 9, pp. 885–896, Sep. 2003.

[33] Online. (2005). "Classic ap profile version 4.03," [Online]. Available: http://www.eugridpma.org/igtf/IGTF-AP-classic-20050905-4-03.pdf

[34] R. Housley, W. Polk, W. Ford, and D. Solo. (2002, Apr.). "Internet X.509 public key infrastructure—Certificate and certificate revocation list (CRL) profile," RFC 3280 (Informational), [Online]. Available: http://www.ietf.org/rfc/rfc3280.txt

[35] Online. (2009) "Health-e-child project," [Online]. Available: http://www.health-e-child.org/links/

[36] C. del Frate, J. Galvez, T. Hauer, D. Manset, R. McClatchey, M. Odeh, D. Rogulin, T. Solomonides, and R. Warren. (2009). "Mammogrid project," [Online]. Available: https://savannah.cern.ch/projects/mammogrid

[37] A. Redolfi, R. McClatchey, A. Anjum, A. Zijdenbos, D. Manset, F. Barkhof, C. Spenger, Y. Legré, L.-O. Wahlund, C. Barattieri, G. B. Frisoni. (2009). "Neugrid project," [Online]. Available: http://www.neugrid.eu/pagine/overview.php

[38] S. Erberich, J. Silverstein, A. Chervenak, R. Schuler, M. Nelson, and C. Kesselman. (2009). "Medicus project," [Online]. Available: http://dev.globus.org/wiki/Incubator/MEDICUS/

[39] R. Sinnott, M. Bayer, A. Stell, and J. Koetsier, "Grid infrastructures for secure access to and use of bioinformatics data: Experiences from the bridges project," in *Proc. Availability, Reliabil. Security, Int. Conf.*, 2006, vol. 0, pp. 950–957.

[40] Online. (2007, Feb.). "SHARE: Technology and security roadmap," [Online]. Available: http://wiki.healthgrid.org/index.php/Share_Roadmap_I

[41] J. Montagnat, Á. Frohner, D. Jouvenot, C. Pera, P. Kunszt, B. Koblitz, N. Santos, C. Loomis, R. Texier, D. Lingrand, P. Guio, R. B. D. Rocha, A. S. de Almeida, and Z. Farkas, "A secure grid medical data manager interfaced to the glite middleware," *J. Grid Comput.*, vol. 6, no. 1, pp. 45–59, 2008.

[42] J. Kubiatowicz, D. Bindel, Y. Chen, S. E. Czerwinski, P. R. Eaton, D. Geels, R. Gummadi, S. C. Rhea, H. Weatherspoon, W. Weimer, C. Wells, and B. Y. Zhao, "Oceanstore: An architecture for global-scale persistent storage," in *Proc. ASPLOS*, 2000, pp. 190–201.

[43] A. Adya, W. J. Bolosky, M. Castro, G. Cermak, R. Chaiken, J. R. Douceur, J. Howell, J. R. Lorch, M. Theimer, and R. Wattenhofer, "Farsite: Federated, available, and reliable storage for an incompletely trusted environment," in *OSDI*, 2002, pp. 1–14.

[44] M. W. Storer, K. M. Greenan, E. L. Miller, and K. Voruganti, "Secure, archival storage with potshards," in *Proc. 5th Conf. USENIX FAST2007 Berkeley*, CA: USENIX Association, pp. 11–11.

[45] Online. (2007). "Cleversafe," [Online]. Available: http://www.cleversafe.com

Jesus Luna received the Engineering Diploma degree from the National Technical Institute, Mexico City, Mexico, in 1995, the M.Sc. degree in computer science from the Technical Institute of Higher Studies of Monterrey, Monterrey, Mexico, in 2003, and the Ph.D. degree in computer architecture from the Technical University of Catalonia, Catalonia, Spain, in 2008, where he graduated with honors.

During 2007–2008, he was a Postdoctoral Researcher with the CoreGRID Network of Excellence, Greece and Cyprus. Since 2009, he has been a Security Researcher with Barcelona Digital Technological Centre, Barcelona, Spain. His current research interests include grid and cloud security, trust management, e-Health privacy, and applied cryptography.

Marios Dikaiakos received the a Dipl.-Ing. (*summa cum laude*) degree from the National Technical University of Athens, Athens, Greece, in 1988, and the M.A. and Ph.D. degrees from Princeton University, Princeton, NJ, in 1991 and 1994, respectively.

He is currently an Associate Professor of computer science at the University of Cyprus, Nicosia, Cyprus, where he is the Leader of the High-Performance Computing Systems Laboratory. His research interests include network-centric computing, with an emphasis on grids, vehicular ad-hoc networks, and the World Wide Web.

Manolis Marazakis received the Ph.D. degree from the Department of Computer Science, University of Crete, Heraklion, Greece, in 2000. His Ph.D. research was focused in the area of service-level agreements in distributed systems.

He is currently an R&D Engineer at Institute of Computer Science, Foundation for Research and Technology—Hellas, Heraklion, where he is engaged in the design, prototyping, and performance evaluation of networked storage systems. From 2001 to 2006, he was an Adjunct Lecturer at the University of Crete, where he taught courses in the area of distributed systems, embedded systems software, and programming. He is also a Founder and a Co-Owner of NovelTech, a software services company offering hosted Web-based application services and custom software project services. His research interests include software layers (firmware, operating systems, and user-level applications) of computer systems as well as the interactions between these layers.

Theodoros Kyprianou, received the M.D. degree from the Medical School, University of Athens, Athens, Greece (full scholarship, 1986–1992) and the Ph.D. degree from the Department of Critical Care Medicine, University of Athens in 2001.

He is board certified in Pulmonary Medicine (1997) and Critical Care Medicine (EDIC 2002). Leading a team of dedicated health professionals, he organized and currently directs the first multidisciplinary, closed-type, Intensive Care Unit in the Republic of Cyprus (Nicosia General Hospital, 2006) as well as the nonprofit organization "Intensive Care Forum" being its Vice-President since 2005. He is currently Vice-President of the "The review bioethics committee for biomedical research on human beings and their biological substances and the clinical trials on medicinal products of human use" (Cyprus National Bioethics Committee) as well as Vice-President of the Cyprus Pulmonology Society. He is also working as a Postdoctoral Scientist at the High-Performance Computing Systems Laboratory, University of Cyprus and teaches human physiology and pathophysiology at the University of Nicosia, Nicosia, Cyprus. His research interests include integration of bio-signals from the critically ill patient, telemedicine, rehabilitation and competency based medical education. .