Future Generation Computer Systems 28 (2012) 583–592

Contents lists available at SciVerse ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Addressing cloud computing security issues

Dimitrios Zissis *, Dimitrios Lekkas

Department of Product and Systems Design Engineering, University of the Aegean, Syros 84100, Greece

ARTICLE INFO

Article history:
Received 14 May 2010
Received in revised form
11 December 2010
Accepted 13 December 2010
Available online 22 December 2010

Keywords:
Cloud computing security
Trusted Third Party
Public key infrastructure
Information and communication security
Trust

ABSTRACT

The recent emergence of cloud computing has drastically altered everyone's perception of infrastructure architectures, software delivery and development models. Projecting as an evolutionary step, following the transition from mainframe computers to client/server deployment models, cloud computing encompasses elements from grid computing, utility computing and autonomic computing, into an innovative deployment architecture. This rapid transition towards the clouds, has fuelled concerns on a critical issue for the success of information systems, communication and information security. From a security perspective, a number of unchartered risks and challenges have been introduced from this relocation to the clouds, deteriorating much of the effectiveness of traditional protection mechanisms. As a result the aim of this paper is twofold; firstly to evaluate cloud security by identifying unique security requirements and secondly to attempt to present a viable solution that eliminates these potential threats. This paper proposes introducing a Trusted Third Party, tasked with assuring specific security characteristics within a cloud environment. The proposed solution calls upon cryptography, specifically Public Key Infrastructure operating in concert with SSO and LDAP, to ensure the authentication, integrity and confidentiality of involved data and communications. The solution, presents a horizontal level of service, available to all implicated entities, that realizes a security mesh, within which essential trust is maintained.

1. Introduction

Throughout computer science history, numerous attempts have been made to disengage users from computer hardware needs, from time-sharing utilities envisioned in the 1960s, network computers of the 1990s, to the commercial grid systems of more recent years. This abstraction is steadily becoming a reality as a number of academic and business leaders in this field of science are spiralling towards cloud computing. Cloud computing is an innovative Information System (IS) architecture, visualized as what may be the future of computing, a driving force demanding from its audience to rethink their understanding of operating systems, client–server architectures, and browsers. Cloud computing has leveraged users from hardware requirements, while reducing overall client side requirements and complexity.

As cloud computing is achieving increased popularity, concerns are being voiced about the security issues introduced through the adoption of this new model. The effectiveness and efficiency of traditional protection mechanisms are being reconsidered, as the characteristics of this innovative deployment model, differ widely from them of traditional architectures. In this paper we attempt to demystify the unique security challenges introduced in a cloud environment and clarify issues from a security perspective. The notion of trust and security is investigated and specific security requirements are documented. This paper proposes a security solution, which leverages clients from the security burden, by trusting a Third Party. The Third Party is tasked with assuring specific security characteristics within a distributed information system, while realizing a trust mesh between involved entities, forming federations of clouds. The research methodology adopted towards achieving this goal, is based on software engineering and information systems design approaches. The basic steps for designing the system architecture include the collection of requirements and the analysis of abstract functional specifications.

2. Grid and cloud computing

Grid Computing emerged in the early 1990s, as high performance computers were inter-connected via fast data communication links, with the aim of supporting complex calculations and data-intensive scientific applications. Grid computing is defined as "a hardware and software infrastructure that provides dependable consistent, pervasive, and inexpensive access to high-end computational capabilities". Cloud Computing has resulted from the convergence of Grid Computing, Utility Computing and SaaS, and essentially represents the increasing trend towards the external deployment of IT resources, such as computational power, storage or business applications, and obtaining them as services [1]. Cloud

* Corresponding author.
E-mail addresses: Dzissis@aegean.gr (D. Zissis), Dlek@aegean.gr (D. Lekkas).

0167-739X/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.future.2010.12.006

computing is a model for enabling convenient, on-demand network access, to a shared pool of configurable computing resources, (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [2].

The name cloud computing, was inspired by the cloud symbol that is often used to represent the Internet in flow charts and diagrams. A distinct migration to the clouds has been taking place over recent years with end users, "bit by bit" maintaining a growing number of personal data, including bookmarks, photographs, music files and much more, on remote servers accessible via a network. Cloud computing is empowered by virtualization technology; a technology that actually dates back to 1967, but for decades was available only on mainframe systems. In its quintessence, a host computer runs an application known as a hypervisor; this creates one or more virtual machines, which simulate physical computers so faithfully, that the simulations can run any software, from operating systems, to end-user applications [3]. At a hardware level, a number of physical devices, including processors, hard drives and network devices, are located in datacenters, independent from geographical location, which are responsible for storage and processing needs. Above this, the combination of software layers, the virtualization layer and the management layer, allow for the effective management of servers. Virtualization is a critical element of cloud implementations and is used to provide the essential cloud characteristics of location independence, resource pooling and rapid elasticity. Differing from traditional network topologies, such as client–server, cloud computing is able to offer robustness and alleviate traffic congestion issues. The management layer is able to monitor traffic and respond to peaks or drops with the creation of new servers or the destruction of non-necessary ones. The management layer has the additional ability of being able to implement security monitoring and rules throughout the cloud. According to Merrill Lynch, what makes cloud computing new and differentiates it from Grid Computing is virtualization: "Cloud computing, unlike grid computing, leverages virtualization to maximize computing power. Virtualization, by separating the logical from the physical, resolves some of the challenges faced by grid computing" [4]. While Grid Computing achieves high utilization through the allocation of multiple servers onto a single task or job, the virtualization of servers in cloud computing achieves high utilization by allowing one server to compute several tasks concurrently [5]. While most authors acknowledge similarities among those two paradigms, the opinions seem to cluster around the statement that cloud computing has evolved from Grid Computing and that Grid Computing is the foundation for cloud computing.

In cloud computing, the available service models are:

- *Infrastructure as a Service (IaaS)*. Provides the consumer with the capability to provision processing, storage, networks, and other fundamental computing resources, and allow the consumer to deploy and run arbitrary software, which can include operating systems and applications. The consumer has control over operating systems, storage, deployed applications, and possibly limited control of select networking components.
- *Platform as a Service (PaaS)*. Provides the consumer with the capability to deploy onto the cloud infrastructure, consumer-created or acquired applications, produced using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.
- *Software as a Service (SaaS)*. Provides the consumer with the capability to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices, through a thin client interface, such as a web browser (e.g. web-based e-mail). The consumer does not manage or control the underlying cloud infrastructure, including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

Four deployment models have been identified for cloud architecture solutions, described below:

- *Private cloud*. The cloud infrastructure is operated for a private organization. It may be managed by the organization or a third party, and may exist on premise or off premise.
- *Community cloud*. The cloud infrastructure is shared by several organizations and supports a specific community that has communal concerns (e.g., mission, security requirements, policy, and compliance considerations). It maybe managed by the organizations or a third party, and may exist on premise or off premise.
- *Public cloud*. The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.
- *Hybrid cloud*. The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology, that enables data and application portability (e.g., cloud bursting for load-balancing between clouds) [2].

Cloud computing is viewed as one of the most promising technologies in computing today, inherently able to address a number of issues. A number of key characteristics of cloud computing have been identified [6,7]:

Flexibility/Elasticity: users can rapidly provision computing resources, as needed, without human interaction. Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out or up.

Scalability of infrastructure: new nodes can be added or dropped from the network as can physical servers, with limited modifications to infrastructure set up and software. Cloud architecture can scale horizontally or vertically, according to demand.

Broad network access. Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous platforms (e.g., mobile phones, laptops, and PDAs).

Location independence. There is a sense of location independence, in that the customer generally has no control or knowledge over the exact location of the provided resources, but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter).

Reliability improves through the use of multiple redundant sites, which makes cloud computing suitable for business continuity and disaster recovery.

Economies of scale and cost effectiveness. Cloud implementations, regardless of the deployment model, tend to be as large as possible in order to take advantage of economies of scale. Large cloud deployments can often be located close to cheap power stations and in low-priced real estate, to lower costs.

Sustainability comes about through improved resource utilization, more efficient systems, and carbon neutrality.

Cloud implementations often contain advanced security technologies, mostly available due to the centralization of data and universal architecture. The homogeneous resource pooled nature of the cloud, enables cloud providers, to focus all their security resources on securing the cloud architecture. At the same time, the automation capabilities within a cloud, combined with the large focused security resources, usually result in advanced security capabilities. Maintaining a perspicacious vision is essential in a field

that is evolving exponentially. Cloud computing is not a panacea and many believe it to be a market-driven hype. Cautiousness is necessary, so as to not be carried away by the caprice of the moment. Cloud computing in its quintessence, has the capability to address a number of identified deficiencies of traditional architectures due to its unique characteristics, but the adoption of this innovative architecture may introduce a number of additional uncategorized threats (Fig. 1).

3. Cloud computing security

3.1. Trust

Trust is not a new research topic in computer science, spanning areas as diverse as security and access control in computer networks, reliability in distributed systems, game theory and agent systems, and policies for decision making under uncertainty [8]. Perhaps the most notable example was the development of the Trusted Computer System Evaluation Criteria (TCSEC) [9] in the late 70s and early 80s. Here, trust was used in the process of convincing observers that a system (model, design or implementation) was correct and secure [10].

The concept of trust, adjusted to the case of two parties involved in a transaction, can be described as follows: ''An entity A is considered to trust another entity B when entity A believes that entity B will behave exactly as expected and required'' [11]. Thereinafter, an entity can be considered trustworthy, if the parties or people involved in transactions with that entity rely on its credibility. In general, the concept described above can be verbally represented by the term reliability, which refers to the quality of a person or entity that is worthy of trust. Trust in the information society is built on various different grounds, based on calculus, on knowledge or on social reasons [12]. The notion of trust in an organization could be defined as the customer's certainty that the organization is capable of providing the required services accurately and infallibly. A certainty which also expresses the customer's faith in its moral integrity, in the soundness of its operation, in the effectiveness of its security mechanisms, in its expertise and in its abidance by all regulations and laws, while at the same time, it also contains the acknowledgement of a minimum risk factor, by the relying party. The notion of security refers to a given situation where all possible risks are either eliminated or brought to an absolute minimum [13].

Trust in a cloud environment depends heavily on the selected deployment model, as governance of data and applications is outsourced and delegated out of the owner's strict control. In traditional architectures, trust was enforced by an efficient security policy, which addressed constraints on functions and flow among them, constraints on access by external systems and adversaries including programs and access to data by people. In a cloud deployment, this perception is totally obscured. In the case of public or community clouds, control is delegated to the organization owning the infrastructure. When deploying on a public cloud, control is mitigated to the infrastructure owner to enforce a sufficient security policy that guarantees that appropriate security activities are being performed to ensure that risk is reduced. This introduces a number of risks and threats, as essentially security is related to trusting the processes and computing base implemented by the cloud owner. It is crucial to differentiate between deployment models, as a private cloud, where the infrastructure is operated and managed on premise by a private organization, does not introduce additional unique security challenges, as trust remains within the organization. In such a situation the infrastructures owner remains the data and process owner.

Most importantly the cloud environment deteriorates the perception of perimeter security. Perimeter security is a set of physical and programmatic security policies that provide levels of protection on a conceptual borderline against remote malicious activity. Traditionally, it is believed that any connectivity to systems or organizations outside of an organization provides an opening for unauthorized entities (personnel or processes) to gain access or tamper with information resources. Upon this static conceptual boundary, security controls were deployed to protect the Information System within it. In a cloud computing model, the perimeter becomes fuzzy, weakening the effectiveness of this measure. The emergence of cloud service models, is expected to lead to a deconstruction of the application services as they are already delivered in existing ''closed'' service provisioning environments [14]. From the traditional viewpoint of perimeter security, the cloud appears outside the trust borderline and should be viewed with suspicion, but this adversely leads to not trusting essential business processes and services that have been outsourced. It has become impossible to place a virtual moat around an organizations castle, as an abundance of services have been outsourced. The ability to clearly identify, authenticate, authorize and monitor who or what is accessing the assets of an organization is essential to protecting an IS from threats and vulnerabilities. Separation is the key ingredient of any secure system, and is based on the ability to create boundaries between those entities that must be protected and those which cannot be trusted [15].

This paper proposes using a Trusted Third Party within a cloud environment by enabling trust and using cryptography to ensure the confidentiality, integrity and authenticity of data and communications, while attempting to address specific security vulnerabilities. The notion of trust against a Third Party, expresses the customer's faith in specific operational, ethical and quality characteristics, while it also includes the acknowledgement of a minimum risk factor. The relying party customers trust the TTP for the security support it is supposed to offer in all transactions [16]. The scope of a TTP within an Information System is to provide end-to-end security services, which are scalable, based on standards and useful across different domains, geographical areas and specialisation sectors. Introducing a trusted third party can specifically address the loss of the traditional security boundary by producing trusted security domains and enabling cooperation between these. TTP is an ideal security facilitator in a distributed cloud environment where entities belonging to separate administrative domains, with no prior knowledge of each other, are required to establish secure interactions. A TTP is essentially a Trusted Authority delegated with the responsibility of addressing a number of security issues in a multilevel distributed environment.

3.2. Security identification of threats

Essentially securing an Information System (IS), involves identifying unique threats and challenges which need to be addressed by implementing the appropriate countermeasures. Ultimately, the identified security requirements and selected security controls are introduced to the standard systems engineering process, to effectively integrate the security controls with the information systems functional and operational requirements, as well as other pertinent system requirements (e.g., reliability, maintainability, supportability) [17]. Cloud computing due to its architectural design and characteristics imposes a number of security benefits, which include centralization of security, data and process segmentation, redundancy and high availability. While many traditional risks are countered effectively, due to the infrastructures singular characteristics, a number of distinctive security challenges are introduced. Cloud computing has ''unique attributes that require risk assessment in

areas such as availability and reliability issues, data integrity, recovery, and privacy and auditing" [18].

Security in general, is related to the important aspects of confidentiality, integrity and availability; they thus become building blocks to be used in designing secure systems. These important aspects of security, apply to the three broad categories of assets which are necessary to be secured, data, software and hardware resources. The cloud infrastructure proposes unique security challenges which need to be considered in detail.

3.2.1. Confidentiality and privacy

Confidentiality refers to only authorized parties or systems having the ability to access protected data. The threat of data compromise increases in the cloud, due to the increased number of parties, devices and applications involved, that leads to an increase in the number of points of access. Delegating data control to the cloud, inversely leads to an increase in the risk of data compromise, as the data becomes accessible to an augmented number of parties. A number of concerns emerge regarding the issues of multitenancy, data remanence, application security and privacy [19].

Multitenancy refers to the cloud characteristic of resource sharing. Several aspects of the IS are shared including, memory, programs, networks and data. Cloud computing is based on a business model in which resources are shared (i.e., multiple users use the same resource) at the network level, host level, and application level. Although users are isolated at a virtual level, hardware is not separated. With a multitenant architecture, a software application is designed to virtually partition its data and configuration so that each client organization works with a customized virtual application instance. Multitenancy, is relative to multitasking in operating systems. In computing, multitasking is a method by which multiple tasks, also known as processes, share common processing resources such as a CPU. Multitenancy, as multitasking, presents a number of privacy and confidentiality threats. Object reusability is an important characteristic of cloud infrastructures, but reusable objects must be carefully controlled lest they create a serious vulnerability. Data confidentiality could be breached unintentionally, due to data remanence. Data remanence is the residual representation of data that have been in some way nominally erased or removed. Due to virtual separation of logical drives and lack of hardware separation between multiple users on a single infrastructure, data remanence may lead to the unwilling disclosure of private data. But also maliciously, a user may claim a large amount of disk space and then scavenge for sensitive data.

Data confidentiality in the cloud is correlated to user authentication. Protecting a user's account from theft is an instance of a larger problem of controlling access to objects, including memory, devices, software etc. Electronic authentication is the process of establishing confidence in user identities, electronically presented to an information system. Lack of strong authentication can lead to unauthorized access to users account on a cloud, leading to a breach in privacy.

Software confidentiality is as important as data confidentiality to the overall system security. Software confidentiality refers to trusting that specific applications or processes will maintain and handle the user's personal data in a secure manner. In a cloud environment the user is required to delegate "trust" to applications provided by the organization owning the infrastructure. Software applications interacting with the user's data must be certified not to introduce additional confidentiality and privacy risks. Unauthorized access can become possible through the exploitation of an application vulnerability or lack of strong identification, bringing up issues of data confidentiality and privacy. In addition, the cloud provider is responsible for providing secure cloud instances, which should ensure users privacy.

Privacy is the desire of a person to control the disclosure of personal information. Organizations dealing with personal data are required to obey to a country's legal framework that ensures appropriate *privacy* and confidentiality protection. The cloud presents a number of legal challenges towards privacy issues involved in data stored in multiple locations in the cloud, additionally increasing the risk of confidentiality and privacy breaches. Instead of its data being stored on the company's servers, data is stored on the service provider's servers, which could be in Europe, Asia, or anywhere else. This tenet of cloud computing conflicts with various legal requirements, such as the European laws that require that an organization know where the personal data in its possession is at all times.

3.2.2. Integrity

A key aspect of Information Security is integrity. Integrity means that assets can be modified only by authorized parties or in authorized ways and refers to data, software and hardware. *Data Integrity* refers to protecting data from unauthorized deletion, modification or fabrication. Managing an entity's admittance and rights to specific enterprise resources ensures that valuable data and services are not abused, misappropriated or stolen. By preventing unauthorized access, organizations can achieve greater confidence in data and system integrity. Additionally, such mechanisms offer the greater visibility into determining who or what may have altered data or system information, potentially affecting their integrity (accountability). Authorization is the mechanism by which a system determines what level of access a particular authenticated user should have to secured resources controlled by the system. Due to the increased number of entities and access points in a cloud environment, authorization is crucial in assuring that only authorized entities can interact with data.

A cloud computing provider is trusted to maintain data integrity and accuracy. The cloud model presents a number of threats including sophisticated insider attacks on these data attributes.

Software Integrity refers to protecting software from unauthorized deletion, modification, theft or fabrication. Deletion, modification or fabrication can be intentional or unintentional. For instance a disgruntled employee may intentionally modify a program to fail when certain conditions are met or when a certain time is reached. Cloud computing providers implement a set of software interfaces or APIs that customers use to manage and interact with cloud services. In addition to previously mentioned threats, the security of cloud services depends heavily on the security of these interfaces as an unauthorized user gaining control of them could alter delete or fabricate user data. In the cloud, responsibility for the protection of the software's integrity is transferred to the software's owner or administrator. Hardware and network integrity is an additional issue that needs to be addressed by the cloud provider, as he is burdened with protecting the underlying hardware from theft, modification and fabrication.

3.2.3. Availability

Availability refers to the property of a system being accessible and usable upon demand by an authorized entity. System availability includes a systems ability to carry on operations even when some authorities misbehave. The system must have the ability to continue operations even in the possibility of a security breach. Availability refers to data, software but also hardware being available to authorized users upon demand. Leveraging users from hardware infrastructure demands, generates a heavy reliance on the ubiquitous network's availability. The network in now burdened with data retrieval and processing. The cloud owner needs

Table 1
User-specific security requirements.

Level	Service level	Users	Security requirements	Threats
Application level	Software as a Service (SaS)	End client applies to a person or organization who subscribes to a service offered by a cloud provider and is accountable for its use	• Privacy in multitenant environment • Data protection from exposure (remnants) • Access control • Communication protection • Software security • Service availability	• Interception • Modification of data at rest and in transit • Data interruption (deletion) • Privacy breach • Impersonation • Session hijacking • Traffic flow analysis • Exposure in network
Virtual level	Platform as a Service (PaS) Infrastructure as a Service (IaS)	Developer–moderator applies to a person or organization that deploys software on a cloud infrastructure	• Access control • Application security • Data security, (data in transit, data at rest, remanence) • Cloud management control security • Secure images • Virtual cloud protection • Communication security	• Programming flaws • Software modification • Software interruption (deletion) • Impersonation • Session hijacking • Traffic flow analysis • Exposure in network • Defacement • Connection flooding • DDOS • Impersonation • Disrupting communications
Physical level	Physical datacenter	Owner applies to a person or organization that owns the infrastructure upon which clouds are deployed	• Legal not abusive use of cloud computing • Hardware security • Hardware reliability • Network protection • Network resources protection	• Network attacks • Connection flooding • DDOS • Hardware interruption • Hardware theft • Hardware modification • Misuse of infrastructure • Natural disasters

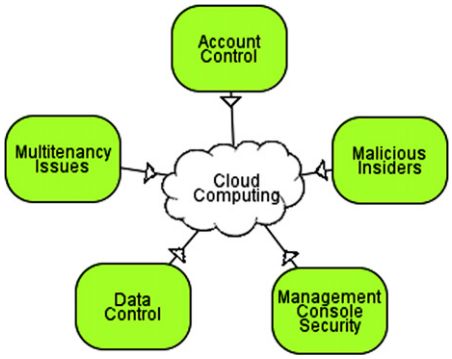

Fig. 1. Categorization of threats.

to guarantee that information and information processing is available to clients upon demand. System availability includes a systems ability to carry on operations even when some authorities misbehave. The system must have the ability to continue operations even in the possibility of a security breach. Cloud computing services present a heavy reliance on the resource infrastructures and network availability at all times.

Understanding and clearly documenting specific user requirements is imperative in designing a solution targeting at assuring these necessities. Verifying identities many of which share common fundamental security requirements, and determining specific needs for data protection and information security can be one of the most complex elements of IS design. This multiuser distributed environment proposes unique security challenges, dependent on the level at which the user operates, application, virtual or physical (Table 1).

The security objectives within a distributed system are essentially [20]:

- to ensure the availability of information communicated between or held within participating systems;
- to maintain the integrity of information communicated between or held within participating systems, i.e. preventing the loss or modification of information due to unauthorized access, component failure or other errors;
- to maintain the integrity of the services provided, i.e. confidentiality and correct operation;
- to provide control over access to services or their components to ensure that users may only USC services for which they are authorized;
- to authenticate the identity of communicating partners (peer entities) and where necessary (e.g. for banking purposes) to ensure non-repudiation of data origin and delivery; and
- where appropriate, to provide secure interworking with the non-open systems world.

While adding,

- To ensure the confidentiality of information held on participating systems.
- Clear separation of data and processes on the virtual level of the cloud, ensuring zero data leakage between different applications.
- To maintain the same level of security when adding or removing resources on the physical level.

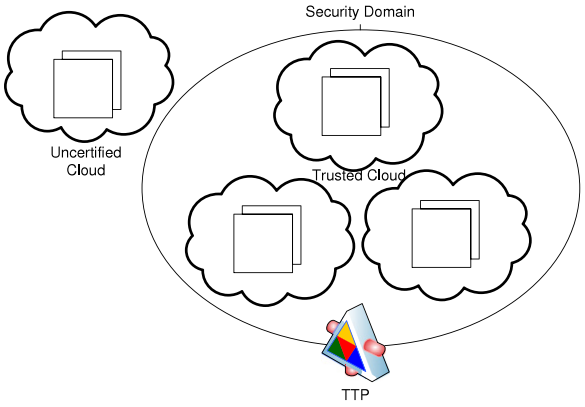

Fig. 2. TTP to enable cloud federations.

4. Trusted Third Party

We claim that employing Trusted Third Party services within the cloud, leads to the establishment of the necessary Trust level and provides ideal solutions to preserve the confidentiality, integrity and authenticity of data and communications [21] (Fig. 2). In cryptography, a Trusted Third Party (TTP) is an entity which facilitates secure interactions between two parties who both trust this third party. The scope of a TTP within an Information System is to provide end-to-end security services, which are scalable, based on standards and useful across different domains, geographical areas and specialisation sectors. The establishment and the assurance of a trust relationship between two transacting parties shall be concluded as a result of specific acceptances, techniques and mechanisms. The Third Party reviews all critical transaction communications between the parties, based on the ease of creating fraudulent digital content. Introducing a Trusted Third Party can specifically address the loss of the traditional security boundary by producing trusted security domains. As described by Castell, "A Trusted Third Party is an impartial organization delivering business confidence, through commercial and technical security features, to an electronic transaction. It supplies technically and legally reliable means of carrying out, facilitating, producing independent evidence about and/or arbitrating on an electronic transaction. Its services are provided and underwritten by technical, legal, financial and/or structural means" [22].

This infrastructure leverages a system of digital certificate distribution and a mechanism for associating these certificates with known origin and target sites at each participating server. TTP services are provided and underwritten not only by technical, but also by legal, financial, and structural means [22,23]. TTPs are operationally connected through chains of trust (usually called certificate paths) in order to provide a web of trust forming the notion of a Public Key Infrastructure (PKI). Public Key Infrastructure provides technically sound and legally acceptable means to implement:

- Strong Authentication: The control of authenticity, the process of identification of parts involved in electronic transactions or exchange of information with electronic means.
- Authorization: The authenticated access to resources, database and informative systems, according to the user's permission rights and the roles
- Data Confidentiality: The protection of information either locally stored or in transmission from unauthorized access.
- Data Integrity: The protection of information either locally stored or in transmission from unauthorized modification.
- Non-Repudiation: Ensuring that no part of an electronic transaction can deny its attendance in it.

PKI in a distributed information system, benefits from the coupling with a directory. A directory is a set of objects with similar attributes organized in a logical and hierarchical manner. The lightweight directory access protocol, or LDAP, is the Internet standard way of accessing directory services that conform to the X.500 data model. LDAP has become the predominant protocol in support of PKIs accessing directory services for certificates and certificate revocation lists (CRLs) and is often used by other (web) services for authentication.

A directory when coupled with PKI can be used to distribute [24]:

- Certificates, for applications such as e-mail in which an end-user certificate must be obtained before an encrypted message is sent.
- Certificate status information, such as certificate revocation lists (CRLs).
- Private keys, when portability is required in environments where users do not use the same machine each day. The directory stores encrypted private keys which are decrypted at the remote workstation using a password provided by the user.

PKI deployed in concert with Single-Sign-On (SSO) mechanisms are ideal for distributed environments, such as cloud environments, where users navigate between an abundance of cross-organization boundaries. In a Single-Sign-On environment, a user does not need to repeatedly enter passwords to access resources across a network. Instead the user signs on once using a password, smart card, or other authentication mechanism, and thereby obtains access to multiple resources on different machines. PKI-based Single-Sign-On mechanisms are indispensable within a cloud environment, since they provide the means for a smooth, transparent strong authentication across different physical resources. SSO in concert with PKI enhances complex free, authorization and authentication processes. In practice this results in enhancing the security of the whole infrastructure, among other evident technical issues, because a sufficient level of usability is assured.

The trusted third party can be relied upon for:

- Low and High level confidentiality.
- Server and Client Authentication.
- Creation of Security Domains.
- Cryptographic Separation of Data.
- Certificate-Based Authorization.

4.1. Low and high level confidentiality

Securing data travelling over the network is a hard and highly complex issue, while the threat of data modification and data interruption is continuously rising. A cloud environment increases this complexity as it does not only require protection of traffic towards the cloud but additionally between cloud hosts, as they lack a traditional physical connection. PKI enables implementing IPSec or SSL for secure communications.

IPSec is an IP layer protocol that enables the sending and receiving of cryptographically protected packets of any kind (TCP, UDP, ICMP, etc.) without any modification. IPSec provides two kinds of cryptographic services. Based on necessity, IPSec can provide confidentiality and authenticity, or it can provide authenticity only [25]. IPsec users are able to authenticate themselves to the peer entity, using PKI certificates in a way that enhances scalability, because only the trusted CA certificate(s) need to be transmitted beforehand. SSL protocol generates end-to-end encryption by interfacing between applications and the TCPIP protocols to provide client–server authentication and an encrypted communications channel between client–server.

Due to the cloud environments unique characteristics, communications are required to be protected between users and hosts but

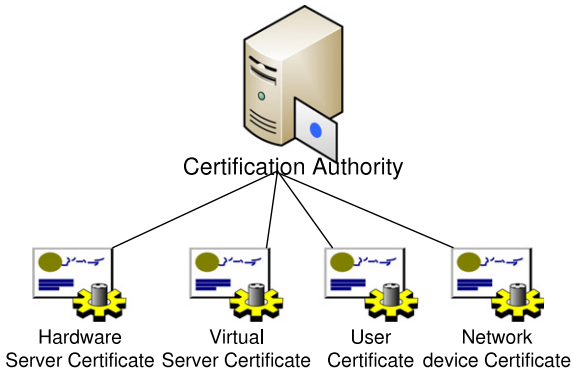

Fig. 3. Certificate categories.

also from host-to-host. Choosing IPSec or SSL depends on the diverse needs and security requirements. IPSec is compatible with any application but requires an IPSec client to be installed on each remote device (PC, PDA, etc.) to add the encryption. In contrast, SSL is built into every browser, so no special client software is required. As the cloud environment promotes use by heterogeneous platforms it is unacceptable to require users to install an IPSec client for encryption. In addition as cloud services are mostly accessed through browsers, SSL has many benefits for client to host communications. On the other hand, IPSec supports using compression making it a more efficient choice for host-to-host communications. This paper proposes implementing IPSec for encrypting communications for host-to-host communications and SSL for Client-to-Cloud communications.

4.2. Server and client authentication

In a cloud environment a Certification authority is required to certify entities involved in interactions, these include certifying physical infrastructure servers, virtual servers, environments users and the networks devices (Fig. 3). The PKI certification authority is responsible for generating these required certificates while registering these within the trust mesh. In other words, a Certification Authority builds the necessary strong credentials for all the physical or virtual entities involved in a cloud and it therefore builds a security domain with specific boundaries within the otherwise fuzzy set of entities of a cloud.

Digital signatures in combination with SSO and Ldap, implement the strongest available authentication process in distributed environments while guaranteeing user mobility and flexibility. The signing private key can be used to authenticate the user automatically and transparently to other servers and devices around the network whenever he/she wants to establish a connection with them.

While the cloud is becoming the common operating platform, every service is going to require a secure authentication and authorization process. As the conceptual boundary between an organizations own service's and outsourced services becomes "fuzzy", the need to adopt Single-Sign-On solution is critical. Users require to make use of applications deployed on their virtual "office" without having to repeat the authentication process on each service (application) provider or maintain numerous passwords, but make use of a single strong authentication process that authorizes them to use services across trusted parties. "*Eight years ago, it was all about securing applications within the enterprise through identity management. Today we talk about securing applications in the cloud with identities originating within the enterprise*" [26].

Shibboleth is standards-based, open source middleware software which provides Web Single Sign On (SSO) across or within

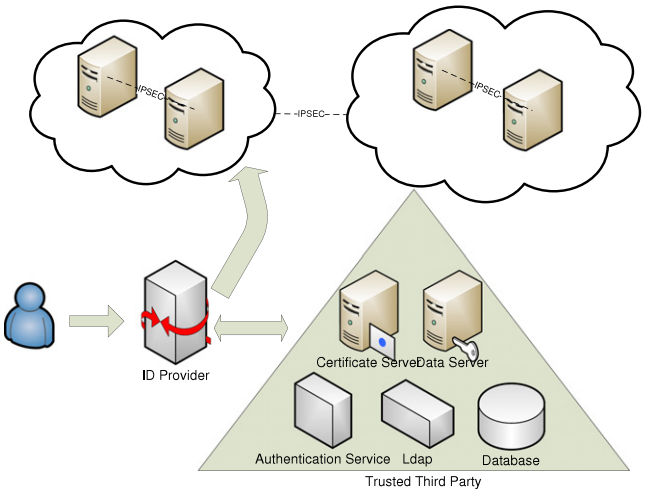

Fig. 4. Authentication in the trusted environment.

organizational boundaries. It allows sites to make informed authorization decisions for individual access of protected online resources in a privacy-preserving manner [27]. Shibboleth technology relies on a third party to provide the information about a user, named attributes. In the proposed system architecture, this is performed by the TTP LDAP repository. It is essential to distinguish the authentication process from the authorization process. During the authentication process a user is required to navigate to his home organization and authenticate himself. During this phase information is exchanged between the user and his home organization only. After the successful authentication of a user, according to the user attributes/credentials, permission to access resources is either granted or rejected. The process in which the user exchanges his attributes with the resource server is the authorization process during which no personal information is leaked and can only be performed after successful authentication (Fig. 4).

To maximize interoperability between communicating parties, it is a necessity to adopt widely used standards. Security Assertion Markup Language (SAML), is an XML-based standard for exchanging authentication and authorization of data between security domains. The primary function of the Shibboleth system is to support identity federation between multiple sites using the SAML protocol standard. The Shibboleth and SAML design processes have been coupled to insure that Shibboleth is standards-based [28]. Because of this design, on a software level, a major part of the Shibboleth system is the OpenSAML libraries, which are also widely used. Both the OpenSAML libraries and the Shibboleth software are developed by the Shibboleth team and released as open source. Shibboleth's added value lies in support for privacy, business process improvement via user attributes, extensive policy controls, and large-scale federation support via metadata (Fig. 5).

4.3. Creation of security domains

Introducing federations, in association with PKI and Ldap technology, will leads to efficient trust relationships between involved entities. A federation is a group of legal entities that share a set of agreed policies and rules for access to online resources [29]. A federation provides a structure and a legal framework that enables authentication and authorization across different organizations. Cloud infrastructures can be organized in distinctive security domains (an application or collection of applications that all trust a common security token for authentication, authorization or session management) enabling "Federated clouds". Federated Clouds are a collection of single Clouds that can interoperate, i.e. exchange

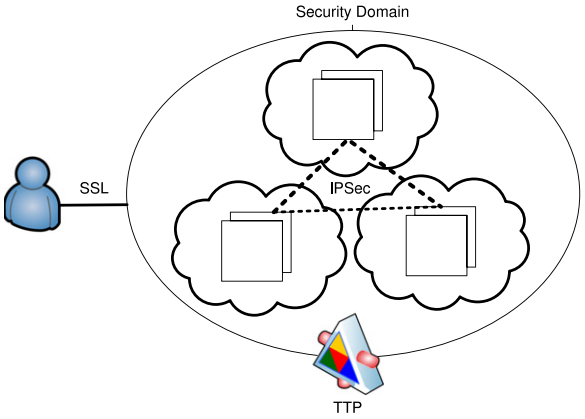

Fig. 5. Security domains.

data and computing resources through defined interfaces. According to basic federation principles, in a Federation of Clouds each single Cloud remains independent, but can interoperate with other Clouds in the federation through standardized interfaces. A federation provides a structure and a legal framework that enables authentication and authorization across different organizations [1].

4.4. Cryptographic separation of data

The protection of personal information or/and sensitive data, within the framework of a cloud environment, constitutes a crucial factor for the successful deployment of SaS and AaS models. Cryptographic Separation in which processes, computations and data are concealed in such a way that they appear intangible to outsiders [30]. Confidentiality and integrity, but also privacy of data can be protected through encryption. Using a combination of asymmetric and symmetric cryptographic (often referred to as hybrid cryptography) can offer the efficiency of symmetric cryptography while maintaining the security of asymmetric cryptography.

4.5. Certificate-based authorization

A cloud environment is a virtual net of several independent domains. In a cloud environment, the relationship between resources and users is more ad hoc and dynamic, resource providers and users are not in the same security domain, and users are usually identified by their characteristics or attributes rather than predefined identities. Therefore, the traditional identity-based access control models are not effective, and access decisions need to be made based on attributes [31]. Certificates issued by a PKI facility can be used for enforcing access control in the Web environment. An example is the use of an extended X.509 certificate that carries role information about a user. These certificates are issued by a certification authority that acts as a trust center in the global Web environment [32]. Attribute certificates contain an attribute–value pair and the principal to whom it applies. They are signed by attribute authorities that have been specified in a use-condition certificate. Attribute based access control, making access decisions based on the attributes of requestors, resources, and the environment, provides the flexibility and scalability that are essential to large-scale distributed systems such as the cloud.

5. Assessment

This paper attempts to propose a security solution to a number of challenges in a cloud environment, which leverages clients from the security burden, by trusting a Third Party. Trust essentially operates in a top-down fashion, as every layer needs to trust the layer immediately below it, and requires a security guarantee at an operational, technical, procedural and legal level to enable secure communications with it (Fig. 6). A trusted certificate serves as a reliable electronic "passport" that establishes an entity's identity, credentials and responsibilities. Trust can be viewed as a chain from the end user, to the application owner, who in turn trusts the infrastructure provider (either at a virtual or hardware level according to the selected service model). A Trusted Third Party is able to provide the required trust by guaranteeing that communicating parties are who they claim to be and have been scrutinized to adhere to strict requirements. This process is performed through the certification process, during which an entity requiring certification is required to conform with a set of policies and requirements. TTP is an ideal security facilitator in a distributed cloud environment where entities belonging to separate administrative domains, with no prior knowledge of each other, require to establish secure interactions.

An end user is required to use his personal digital certificate to strongly authenticate himself with a cloud service and validate his access rights to a required resource. This certificate is used in combination with the service provider's certificate (PaS or IaS level) to create a secure SSL connection between them, thus encrypting exchanged data and guaranteeing their security through the cloud infrastructure (Fig. 7). The user is able to encrypt all personal data stored on the cloud to counter previously identified confidentiality risks. As cloud infrastructure's host a number of services, several applications can be mounted on a virtual server, each requiring separate digital certificates for SSL communications (different ports can be used to support more than one SSL connections to a virtual server).

The application provider can use his own certificate to authenticate himself in communications with the cloud but also use this certificate to encrypt and decrypt application data. These certificates can be enhanced to carry role information about a user or process (extended X.509 certificates). At the lowest level the hardware infrastructure owner makes use of a digital certificate to communicate security between devices and virtual servers but also for authentication purposes if required.

Key management is a critical issue in cloud infrastructures, as the virtualisation of services obscures the identification of the physical key storage location, disabling traditional protection mechanisms. Keys are principally stored and protected at a hardware infrastructure level. In such an environment deploying tamperproof devices for key protection is essential e.g. user smarts cards coupled with Hardware Security Module as part of the virtual deployment.

The proposed solution calls upon cryptography, specifically Public Key Infrastructure, to ensure the authentication, integrity and confidentiality of involved data and communications. A TTP is tasked with assuring specific security characteristics within a cloud environment, while realizing a trust mesh between involved entities, forming federations of clouds. The solution, presents a horizontal level of service, available to all implicated entities, that realizes a security mesh, within which essential trust is maintained. This approach makes use of a combination of Public Key Cryptography, Single-Sign-On technology and LDAP directories to securely identify and authenticate implicated entities. The model presented in this paper offers the advantages of each single technology used and deals with their deficiencies through their combined implementation.

The trusted third party can be relied upon for:

- Low and High level confidentiality.
- Server and Client Authentication.
- Generating Security Domains.

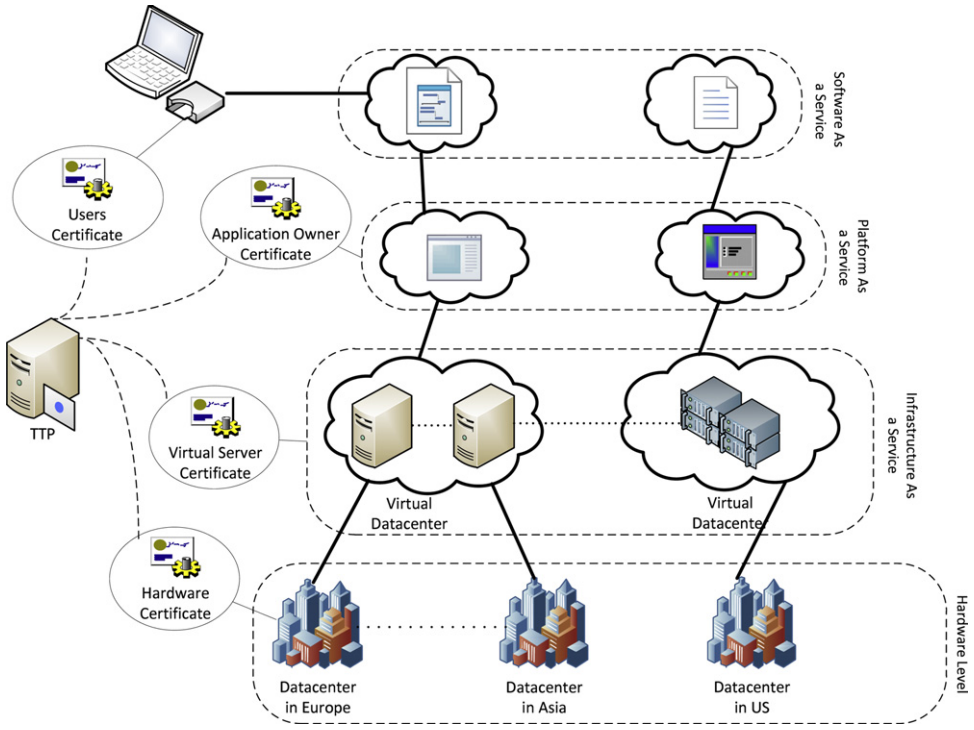

Fig. 6. Trust essentially operates in a top-down fashion, as every layer is required to trust the layer immediately below it.

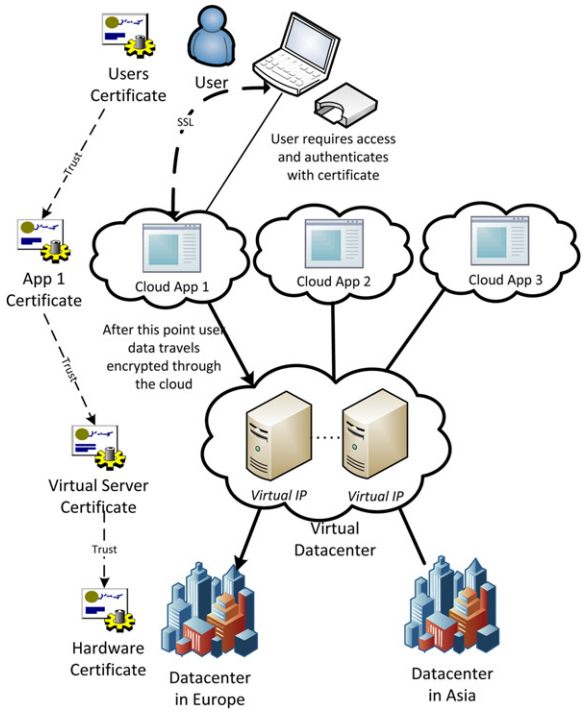

Fig. 7. A user authenticates himself with a cloud service using his personal certificate which in combination with the service provider certificate is used to secure and encrypt all communications.

- Cryptographic Separation of Data.
- Certificate-Based Authorization.

Public key Infrastructure is able to effectively transform security problems into key management issues. Ultimately, the success of the proposed solution, as any cryptographic system, is dependent on controlling access to private keys. An additional important factor as in every centralized system, is system and network performance. Availability, is of amplified importance in a cloud infrastructure, as of the increased performance demands on the network. The Quality of Service provided is a key issue, also in host-to-host communication, as additional encryption processes could deter efficiency. The constant encryption and decryption of data could have a heavy toll on speed, inducing additional processing consumption. Using the cloud infrastructures flexibility within the context of demand on cpu, could leverage the system from this overhead and accelerate encryption/decryption. Currently encryption schemes are being researched that allows data to "searched" without the need of it being decrypted. Future work should focus on improving availability and quality of services provided.

We are currently in the process of researching the development of extended cloud certificates that provide information to end users of the trust path followed on layers below them. These certificates will include extended information on data ownership and responsibilities. These certificates will ensure the authentication, integrity and confidentiality of data but also non-repudiation of transactions at layers much below the user.

6. Conclusion

Inevitably cloud computing will support a surplus of information systems as the benefits outnumber its shortcomings. Cloud computing offers deployment architecture, with the ability to address vulnerabilities recognized in traditional IS but its dynamic characteristics are able to deter the effectiveness of traditional countermeasures. In this paper we have identified generic design principles of a cloud environment which stem from the necessity to control relevant vulnerabilities and threats. To do so, software engineering and information systems design approaches were adopted. Security in a cloud environment requires a systemic point of view, from which security will be constructed on trust, mitigating protection to a trusted third party. A combination of PKI, LDAP and SSO can address most of the identified threats

in cloud computing dealing with the integrity, confidentiality, authenticity and availability of data and communications. The solution, presents a horizontal level of service, available to all implicated entities, that realizes a security mesh through federations, within which essential trust is maintained.

References

[1] K. Stanoevska-Slabeva, T. Wozniak, Grid and Cloud Computing-A Business Perspective on Technology and Applications, Springer-Verlag, Berlin, Heidelberg, 2010.
[2] National Institute of Standards and Technology, The NIST Definition of Cloud Computing, Information Technology Laboratory, 2009.
[3] E. Naone, Technology overview, conjuring clouds, MIT Technology Review, July–August, 2009.
[4] Merrill Lynch, The cloud wars: $100+ billion at stake, Merrill Lynch, 2008.
[5] D. Harris, Why 'grid' doesn't sell, 2008.
[6] G. Reese, Cloud Application Architectures: Building Applications and Infrastructure in the Cloud, in: Theory in Practice, O'Reilly Media, 2009.
[7] B. Rajkumar, C. Yeo, S. Venugopal, S. Malpani, Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility, Future Generation Computer Systems (2009).
[8] D. Artz, Y. Gil, A survey of trust in computer science and the semantic web, Journal of Web Semantics: Science, Services and Agents on the World Wide Web (2007).
[9] DoD Computer Security Center, Trusted computer system evaluation criteria, DoD 5200.28-STD, 1985.
[10] A. Nagarajan, V. Varadharajan, Dynamic trust enhanced security model for trusted platform based, Future Generation Computer Systems (2010) doi:10.1016/j.future.2010.10.008.
[11] International Telecommunication Union, X-509 | ISO/IEC 9594-8, The directory: Public-key and attribute certificate frameworks, ITU, X-Series, 2001.
[12] D. Lekkas, Establishing and managing trust within the public key infrastructure, Computer Communications 26 (16) (2003).
[13] A. Giddens, The Consequences of Modernity, Polity Press, UK, 1991.
[14] K. Tserpes, F. Aisopos, D. Kyriazis, T. Varvarigou, Service selection decision support in the Internet of services, in: Economics of Grids, Clouds, Systems, and Services, in: Lecture Notes in Computer Science, vol. 6296, 2010, pp. 16–33. doi:10.1007/978-3-642-15681-6_2.
[15] R. Sherman, Distributed systems security, Computers & Security 11 (1) (1992).
[16] D. Lekkas, S. Gritzalis, S. Katsikas, Quality assured trusted third parties for deploying secure Internet-based healthcare applications, International Journal of Medical Informatics (2002).
[17] National Institute of Standards and Technology. Guide for mapping types of information and information systems to security categories, NIST 800-60, 2008.
[18] Gartner. Assessing the security risks of cloud computing, Gartner, 2008.
[19] Cloud Security Alliance. Top threats to cloud computing, Cloud Security Alliance, 2010.
[20] R. Sherman, Distributed systems security, Computers & Security 11 (1) (1992).
[21] D. Polemi, Trusted third party services for health care in Europe, Future Generation Computer Systems 14 (1998) 51–59.
[22] S. Castell, Code of practice and management guidelines for trusted third party services, INFOSEC Project Report S2101/02, 1993.
[23] Commission of the European Community. Green paper on the security of information systems, ver. 4.2.1, 1994.
[24] VeriSign. Directories and public—key infrastructure (PKI), Directories and Public—Key Infrastructure, PKI.
[25] A. Alshamsi, T. Saito, A technical comparison of IPSec and SSL, Cryptology (2004).
[26] Cloud Identity Summit, Secure the cloud now, Cloud identity summit, Retrieved on 10/11/2010 from: http://www.cloudidentitysummit.com/.
[27] Internet 2, Shibboleth [Online] 2007, Retrieved on 10/11/2010 from: http://shibboleth.internet2.edu/.
[28] Internet 2, FAQ on SAML and Shibboleth relationship, Shibboleth, Internet 2, 2010. Retrieved on 10/11/2010 from: http://shibboleth.internet2.edu/Shibboleth-SAML-FAQ.html.
[29] UK Federation Information Centre, UK federation information centre, 2007.
[30] C.P. Pfleeger, S.L. Pfleeger, Security in Computing, Prentice Hall, 2002.
[31] B. Lang, I. Foster, F. Siebenlist, R. Ananthakrishnan, T. Freeman, Attribute based access control for grid computing, 2008.
[32] James B.D. Joshi, Walid G. Aref, Arif Ghafoor, Eugene H. Spafford, Security models for web-based applications, Communications of the ACM 44 (2) (2001).

Dimitris Zissis holds a B.Sc. in Computer Science, an M.Sc. in Computing and Information Systems, an MBA in General Management and is currently pursuing a Ph.D. in Information and Communication Security at the University of the Aegean, Department of Product and Systems Design Engineering. He has been involved in a number of EU funded research projects, mostly in the research area of IT Security, involving the development of e-governance solutions and deploying public key infrastructures cryptography.

Dimitrios Lekkas holds a Ph.D. in the area of Information Systems Security, a M.Sc. in Information Technology and a B.Sc. in Mathematics. He is an Assistant Professor at the department of Product and Systems Design Engineering of the University of the Aegean, Greece. He has participated in many research projects funded nationally and by the European Union and published several papers in international journals and conferences. He is a member of the Greek National Educational Network (EDUnet) technical committee and coordinator of the e-School and the e-University Public Key Infrastructure (PKI). His current research interests include design of information infrastructures, computer security, incident response, public key cryptography and digital signatures, database management systems.