User-Controlled Security Mechanism in Data-Centric Clouds

Qin Liu  
College of Computer Science and Electronic Engineering  
Hunan University  
P. R. China, 410082  
Email: gracelq628@hnu.edu.cn

Guojun Wang  
School of Information Science and Engineering  
Central South University  
P. R. China, 410083  
Email: csgjwang@csu.edu.cn

Jie Wu  
Department of Computer and Information Sciences  
Temple University  
Philadelphia, PA 19122, USA  
Email: jiewu@temple.edu

Wei Chang  
Department of Computer and Information Sciences  
Temple University  
Philadelphia, PA 19122, USA  
Email: wei.chang@temple.edu

Abstract—In recent years, cloud computing is no doubt one of the most-talked-about terms in both the industry and academic worlds. In clouds, resources are delivered as services over the Internet in a pay-as-you-go fashion. By leasing cloud platforms to run their business, startups can achieve cost-saving and scale-up elasticity without being concerned about over-provisioning or under-provisioning for a service. Although cloud computing has many benefits, its unique features, such as multi-tenancy and the separation of data administration and data ownership, also raise many security and privacy problems, which have been recognized as the primary concerns hindering clouds’ wide adoption. This paper aims to investigate security and privacy issues in cloud computing, and attempts to identify possible solutions for preserving cloud security. Specifically, we focus on data-centric security, which mainly refers to ensuring data confidentiality, in cloud computing. As an alternative solution to alleviate the risk of data leakage in cloud environments, we provide a user-controlled security mechanism, where the data is depicted with three dimensions on demand, and will be encapsulated in an onion way. The proposed mechanism allows customers to take the initiative to protect their own data. We believe this flexibility could prove to be a major improvement in cloud security if implemented well.

Keywords—cloud computing, data-centric security, user-control security.

I. INTRODUCTION

The term “cloud computing” was first coined by Google CEO Eric Schmidt in 2006, and was immediately popular within industry. Cloud computing makes computing the 5th utility after water, electricity, gas, and telephony [1], and is identified as one of the Top 10 Strategic Technology Trends in 5 successive years [2]. With the great publicity of cloud computing, cloud has been closely related to our daily life. As we are surrounded by various cloud products like cloud players and cloud storages, even those individuals not in the Information Technology (IT) industry have more or less gained some understanding about the buzz word that is cloud.

In clouds, resources (e.g., hardware, software, and data) are delivered as services that can be subscribed and unsubscribed by customers over the Internet in a pay-as-you-go fashion. In other words, everything is a service (XaaS) [3] in clouds, where customers enjoy any desired services on demand, anytime and anywhere, using various kinds of devices connecting to the Internet. Meanwhile, cloud computing, as an evolved paradigm of distributed computing, parallel computing, grid computing, and utility computing, has a lot of merits like fast deployment, pay-for-use, high availability, high scalability, rapid elasticity, low costs, and so on [4]. Especially for startups, the cloud computing paradigm allows them to run their businesses with reduced upfront investment and expected performance, so as to concentrate more on developing the core business without worrying about the underlying deployment details.

Although cloud computing has overwhelming superiority over traditional computing models, the adoption of clouds is still far from expected. The main reason is that customers worry that their sensitive data may be deliberately or unintentionally leaked by the cloud vendors. Actually, the concerns about cloud security are not unnecessary. State-of-art cloud vendors experience noteworthy outages and security breaches from time to time [3]. For example, Gmail’s mass email deletions in 2006, Microsoft Azure had an outage lasting 22 hours in 2008, and the recent news about Apple iCloud leaking out celebrities’ sensitive photos. Gartner [5] indicated that despite of cloud computing’s various benefits, customers should examine carefully potential security risks like privileged user access, regulatory compliance, data location, data segregation, recovery, investigative support, and long-term viability, before selecting a cloud vendor. Therefore, security is still the major concern hindering cloud computing’s further development [6].

IDC estimates that in 2014, the global market for cloud
computing will grow to $100 billion, creating 14 million jobs around the world [7]. However, the essential features of cloud computing not only exacerbate historical security and privacy challenges, but also bring something new. Without understanding its unique security challenges and developing appropriate solutions designed for cloud security, such tremendous potential would be very much in doubt.

The goal of this paper is to investigate security and privacy issues in cloud computing and identify possible solutions for preserving cloud security. Specifically, we focus on data-centric security, which mainly refers to ensuring data confidentiality, in cloud environments. Therefore, we will first introduce the definition and features of cloud computing in Section II, before discussing the security issues in cloud environments in Section III. Then, we will discuss data-centric security problem in cloud computing in Section IV, and propose a possibly feasible approach to provide a secure cloud computing environment in Section V. Finally, we will conclude this paper and discuss the future work in Section VI.

II. BACGROUNDS

Cloud computing originated from industry, but also received wide attention from the academic world. The term “cloud computing” is a hot topic of white papers, academic articles, workshops, conferences, and even magazines. Understanding cloud computing’s definition and principal characteristics is essential to its success. Thus, a large part of published literature is dedicated to framing the definition and characteristics for cloud computing. For example, Vaquero et. al. [8] compared over 20 different cloud definitions from a variety of literatures and drew a definition for cloud as a large pool of easily usable and accessible resources, with features scalability, pay-per-use utility model and virtualization. Buyya et. al. [9] defined a cloud as a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that deliver unified computing resources to customers based on pre-agreed Service Level Agreement (SLA). Armbrust et. al. [10] considered a cloud to include applications as well as the underlying software/hardware that can be delivered as services over the Internet. The illusion of infinite resources, the elimination of up-front investments, and the pay for use pattern are considered as its key characteristics.

Although researchers have tried to define cloud computing, no agreed-upon definition exists yet. Among others, the most authoritative definition comes from the white paper published by the U.S. National Institute of Standards and Technology (NIST) [11]. For the rest of this paper, we will talk about cloud computing in the spirit of the NIST definition, which we believe encompasses the full set of issues of interest.

NIST Definition. Cloud computing is a model for enabling ubiquitous and on-demand access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) over network. This cloud model is composed of three service models, four deployment models, and five essential characteristics.

The NIST cloud definition framework is depicted as shown in Fig. 1. By sharing resources at various levels, cloud computing offers various services, including Software as a Service (SaaS) that allows the cloud customers to control only application configurations, Platform as a Service (PaaS) that allows the cloud customers to control the hosting environments, and Infrastructure as a Service (IaaS) that allows the cloud customers to control everything except the hardware infrastructure. The above service models face different security challenges [12]. IaaS is the foundation of all cloud services, with PaaS built upon it, and SaaS in turn built upon PaaS. Generally speaking, a customer using a service at a lower abstraction level should take responsibility for security in more aspects [13]. For example, in IaaS, the customers are primarily responsible for ensuring the security over infrastructure like the operating system, applications, and so on; in PaaS, customers should protect the applications they build and run on the platforms; in SaaS, the customers have to depend on the cloud vendors for proper security measures.

Moreover, based on different serving objects, the deployment models can be classified into public clouds that are publicly accessible, community clouds that are accessible by several organizations, private clouds that are accessible by a single organization, and hybrid clouds that are a mix of two or more deployment models. From a security perspective, public clouds fully exploit advantages like economies of scale, cost efficiency, etc., but in the mean time face various security challenges: private clouds are more secure than public clouds, but will incur a higher cost consumption and a lower resource utilization; community clouds are trade-off models and hybrid clouds are honored as the most promising ones, which enable greater security and lower up-front investments by storing sensitive data in private clouds and outsourcing publicly available data to public clouds [14].

In this paper, we mainly discuss security challenges and appropriate solutions for public clouds, which face the
greatest security challenges among the above deployment models.

Finally, the five essential characteristics that distinguish cloud computing from other paradigms are as follows:

- **On-demand self-service.** A customer can automatically subscribe to any kinds of services as needed without human interaction with cloud vendors.
- **Broad network access.** A customer can access the subscribed services anytime and anywhere using any devices (e.g., mobile phones, tablets, laptops, and workstations) over networks.
- **Resource pooling.** The cloud vendor’s resources (e.g., storage, processing, and networks) are pooled to serve multiple customers, who are unaware of the exact location of the provided resources.
- **Rapid elasticity.** The cloud vendor seems to have unlimited resources, which can be elastically provisioned and released on demand.
- **Measured service.** Resource usage can be monitored, controlled, and reported, at different abstraction levels.

The emergency of cloud computing has made a tremendous impact on the IT industry over the past few years. Cloud computing employs a service-driven business model, where resources are provided as services in an on-demand basis. For IT companies, cloud computing will bring them a win-win situation. On the one side, the startups like Dropbox and Groupon achieve expected performance with a low cost by running services in the cloud. On the other side, the cloud vendors like Amazon and Google take full advantage of superfluous resources by optimization of resource allocation.

As shown in Fig. 2, cloud vendors can be classified into two types [4]: cloud providers providing underlying raw resources (e.g., storage and processing), and service providers renting cloud providers’ resources to build systems or applications for service provision. For example, GigaVox [15] rents Amazon Elastic Compute Cloud (EC2) to provide video on demand services. Here, GigaVox is the service provider, and Amazon is the cloud provider. The service provider and the cloud provider certainly can be the same entity, e.g., Google Apps resides on Google’s own cloud platform. In this paper, we do not make a distinction between different providers, and collectively refer to them as Cloud Service Providers (CSP). Besides CSPs, the customers, who subscribe the services residing on the cloud infrastructure, or deploy their own applications/systems in clouds, are called end users.

### III. Security Issues in Clouds

As cloud computing achieves increased popularity, concerns are being voiced about the security issues during use. Cloud computing combines known technologies such as distributed computing, parallel computing, virtualization, etc. Therefore, certain well-understood vulnerabilities, such as virtualization vulnerabilities, Distributed Denial of Services (DDoS) attacks, web application vulnerabilities, and so on, also exist in cloud computing (though cloud computing’s essential features make them more significant).

However, due to the unique features of multi-tenancy and remote data store, cloud computing faces the following new security challenges:

- **Multi-tenancy security.** Multi-tenancy is an essential attribute of cloud computing. While the data is moved from a single-tenant to a multi-tenant environment, adversaries may initiate side-channel attacks and covert-channel attacks to gain end users’ private information through traffic patterns and side-channel information [16], [17]. Furthermore, within multi-tenant environments, the fates of end users sharing virtualized resources located on the same host are linked together. One tenant who is a highly targeted attack victim could significantly affect the other tenants. Data separation and VMs’ isolation are essential to preserve multi-tenancy security. A clear boundary for each end user’s data must be ensured at different abstraction levels.

- **Accountability.** To make cloud services accountable and trustworthy, each activity should be binding to its subject [18], [19]. Such bindings must be supported by provable and non-disputable evidences, so that each entity cannot deny what he/she has done. In cloud computing, virtualized resources are delivered as services at different abstraction levels, where the end users and the CSPs play different roles for security. Therefore, while a data breach happens, it is hard to determine which entities should be blamed for it. Furthermore, it is quite common to find a SaaS provider building applications upon a PaaS provider’s platform, which in turn rents infrastructure from an IaaS provider. The existence of “security chain” also exacerbates this situation. Auditability and a well-designed SLA specific to cloud environments may benefit accountability. In this way, each entity must behave according to the predefined SLA. Once a violation is found, the actor should be accountable for its misbehavior without dispute.
• **Bilateral auditing.** In cloud environments, the end users and the CSPs are in different trusted domains, and thus, either party may initiate attacks for their own interest. For example, the malicious end users may use cloud to launch SQL injection attacks, side-channel attacks, or DDoS attacks to compromise the CSP or gain other end users’ private information. For the CSP, it may leak end users’ sensitive data for making profits. Although the SLA stipulates behaviors of the parties concerned, it is hard to effectively perform bilateral auditing to verify whether each party obeys the predefined SLA or not [20]. At the present, it is an open direction for researchers to investigate.

• **Inner attacks.** The clouds as the centralized location of the end users’ data has become a tempting target for cybercrime [21]. Breaking through the cloud environments will potentially attack all the end users’ data. The CSPs’ infrastructure and management capabilities are much more powerful and reliable than those of local machines, but the clouds still face both internal and external security threats. Compared with external attackers, inner attackers pose less threat, but achieves the greater impact [12]. Though CSPs claim that their employees are well-chained and will not peek at end users’ data, service transparency makes it hard for the end users to audit CSPs’ behaviors. Researchers suggest encrypting sensitive data before outsourcing to the clouds [22], due to the high frequency of data breach events.

• **Heterogeneity.** Cloud environments are multi-domain environments in which each domain with different security and privacy requirements employ heterogeneous mechanisms [13]. While data is moved from one CSP (e.g., Amazon EC2 that encrypts data by default) to another (e.g., Microsoft Azure that does not encrypt data), the incompatibility of security policies will cause potential data breaches. To address such heterogeneity issues, standards need to be established across different domains for interoperability, stability, and data security.

The fact that data is shared within the cloud is considered as the core scientific problem that separates cloud computing security from traditional computing [23]. For security, moving data from local machines to cloud platforms is actually a double-edged sword. On the up side, the CSPs have more sophisticated security mechanisms and practices to manage the data, and thus the end users can be alleviated of their responsibility to ensure data security. On the down side, the separation of data administration and data ownership deprives the end users of direct control over their data. In cloud environments, the end users need to rely on third parties that are not fully trusted to make decisions about their data. Therefore, it is critical to have appropriate mechanisms to prevent CSPs from abusing end users’ data without authorization.

**IV. DATA-CENTRIC SECURITY**

As discussed above, cloud computing faces many security challenges, each type of which needs to be treated differently. Due to the limited space, this paper mainly focuses on data-centric security, rather than discussing all of them. According to data ownership, end users can be classified into data owners and data users, as shown in Fig. 3. The data owner uploads his data to clouds maintained by the CSP, and the data user requests data from the CSP after obtaining authorization from the data owner. Data-centric security mainly refers to ensuring the CIA of data in cloud environments.

- **Confidentiality.** Confidentiality refers to the prevention of intentional or unintentional unauthorized disclosure of information. A pivot approach is the encryption of sensitive data before storing. Reinforcing access control, authentication, and authorization can also favor preserving data confidentiality.

- **Integrity.** Integrity needs to ensure that unauthorized modifications are not made to data. Message Authentication Code (MAC) and Digital Signature (DS) are two main approaches to achieve integrity. In a cloud environment, the end users do not locally store any data copies, and thus how to efficiently achieve dynamic, blockless, and stateless verification is of great importance [24], [25].

- **Availability.** Availability needs to ensure the reliable and timely access to data or resources. To provide ubiquitous always-on access, a CSP maintains multiple replicas for each data on distributed servers [26]. A key problem of using the replication technique in cloud computing is that it is very expensive to achieve strong consistency to ensure that a user always sees the latest updates [27].

CIA are the important pillars for ensuring security in either a traditional computing paradigm or cloud computing paradigm. Next, we mainly investigate how to preserve data confidentiality in clouds through cryptography techniques.
For confidentiality, existing researches suggest to store only the encrypted data on untrusted servers [22]. In this way, only the authorized entities can decrypt the data with appropriate keys. The unauthorized entities, even if the CSPs, cannot know data contents. Actually, the state-of-art CSPs already adopt cryptographic techniques to preserve data confidentiality. For example, Amazon EC2 encrypt users’ data by default, and Amazon Simple Storage Service (S3) allows users to encrypt their data before outsourcing.

In this paper, we depict data with three dimensions: object, action, and time, as shown in Fig. 4. The object dimension describes the data users who have rights to access such data, with a default value “public” meaning that this kind of data can be accessed by any data users. The time dimension denotes the length of the access right of the object, with a default value “*” meaning that the object’s access right is valid in the whole life cycle. The action dimension describes the read right, write right, and search right of the object, where the default action is “read”. These dimensions stipulate that only the object can play the action on the data during the time. For example, data D is depicted as \{ “Alice”, “nurses or doctors in hospital A”: {“write”, “read”}, \{ “t1 ∼ t2”, “t3 ∼ t4”}\} means that Alice has the right to read D during time t1 and t2, and nurses or doctors in hospital A can read D during time t3 and t4. For data confidentiality, access to each data should follow its three dimensions.

V. USER-CONTROLLED SECURITY MECHANISM

In cloud computing, the physical security boundary does not exist at all, and traditional security techniques such as physical isolation and access control cannot be applied directly to solve all its security issues. Therefore, improvements on existing solutions as well as more mature and newer solutions are imperative to ensure its further development. We believe that a good design would offer a choice of security level and mechanisms [28]. That is, the users have the ability to customize their desired security level and mechanism on demand.

With the spirit of XaaS, we consider security as a kind of service in clouds, and propose a user-controlled security mechanism, which allows the users to take the initiative to protect the security of their own data. The proposed mechanism is shown in Fig. 5. The data owner first specifies three dimensions for each data before uploading, then encapsulates data according to data dimensions, as shown in Fig. 6. Each data is encrypted with three onions: search onion, write onion, and read onion.

Search onion. The clouds have been building up the capacity to store huge amounts of digital data. To allow users to retrieve only the files of their interests without leaking any information to the CSP, we first associate each piece of data with an index that includes several keywords describing the data content. Then, index is encrypted with the search layer, which can be encapsulated with searchable encryption [29], [30]. With searchable encryption, the service provider can perform keyword-based searches on ciphertexts without knowing user keywords and the file contents. To achieve authorized search, i.e., only the users who satisfy the object dimension can perform searches, the work in [31] and [32] can be applied here.

Read onion. In cloud computing environments, data is generally shared by many data users of different roles and attributes. Attribute-Based Encryption (ABE) [33], [34] can be applied to achieve fine-grained access controls on the encrypted data. In ABE, users are identified by a set of attributes rather than an exact identity. The data is encrypted with an attribute-based access structure, such that only the users whose attributes satisfy the access structure can decrypt the ciphertext using their private keys. Specifically, the read layer may encrypt data content with a symmetric key, which is in turn encrypted with ABE over a specific access structure. Furthermore, we can apply proxy re-encryption (PRE) [35] into ABE for ensuring dynamic access control on ciphertexts.

Write onion. The content can be encrypted with homomorphic encryption [36], where the computations can be performed directly on the ciphertexts without decryption. For example, the Paillier cryptosystem [37] allows for the direct performance of multiplication and exponentiation operations on ciphertext. Therefore, the data users can update data contents directly without decapsulating the write onion.

After uploading the above onions to the cloud, the data
The user-controlled security mechanism. 

The user owner authorizes the data users with different access rights, including write, read, and search, for his/her encrypted data. As receiving a data request from a user, the CSP will re-encapsulate each onion with a time layer. Therefore, an authorized user can perform a specific action on the data only when both the time dimension and object dimension are matched.

VI. CONCLUSION

Despite a bit of hype, cloud computing is undeniably a fundamental trend of IT technologies. More and more IT companies are diving into launching cloud products, such as Amazon’s EC2, Google’s AppEng, Microsoft’s Azure, etc. However, security is the main obstacle hindering the wide adoption of cloud computing. In this paper, we investigate definitions, essential characteristics, and the security and privacy challenges in cloud environments. We argue that, to bridge the gap between its mature business model and immature security mechanisms, existing security and privacy solutions should be reevaluated and improved with regard to their applicability in cloud computing. As a possible feasible solution, we propose a user-defined security mechanism, which allows the users to define a three dimensions for each data, and encapsulates data in an onion way, before uploading. We envision that, the user-defined security, which allows the users to take the initiative to protect their own data, is a promising solution for data security in cloud computing.

ACKNOWLEDGMENTS

This work was supported in part by NSFC grants 61272151, 61472451, 61402161, by ISTCP grant 2013DF-B10070, by the China Hunan Provincial Science Technology Program under Grant 2012GK4106, by the Ministry of Education Fund for Doctoral Disciplines in Higher Education under Grant 20111062110043, and by the Fundamental Research Funds for the Central Universities: NSF grants CNS 149860, CNS 1461932, CNS 1460971, CNS 1439672, CNS 1301774, ECCS 1231461, ECCS 1128209, and CNS 1138963.

REFERENCES


[7] “Idc predicts 2014 will be a year of escalation, consolidation, and innovation as the transition to it’s “3rd platform” accelerates,” http://www.idc.com/getdoc.jsp?containerId=prUS24472713.

