Essential Cloud Storage Systems Security Methodologies

Mervat Hashem
College of Information Science and Engineering
Hunan University
Changsha, 410082,
China

Zhiyong Li
College of Information Science and Engineering
Hunan University
Changsha, 410082,
China

Alaa AlGhamry
Arab Open University
Saudi Arabia, Jeddah, KSA

ABSTRACT
The demand to have secure environments for systems' information on cloud systems is arising since this should be more secure and far away from our systems damage disasters and data corruption. The dream became real with the cloud computing. Cloud computing is being more famous and used especially in the enterprises where the cost matters. These enterprises can reduce IT costs by using cloud computing services. Most of the enterprises that don't use cloud computing have a fear of the platform security and the information security. The users always seek for verifying the confidentiality and integrity of their data and computations before using cloud computing. Many researches have been established in these two fields for the infrastructure, platform security, data confidentiality and integrity. This paper, presents two examples for platform security and a comparison between three techniques for how can users secure the data and computations, at cloud environment, to keep the confidentiality and integrity. The ultimate goal of providing such a comparison is to encourage enterprises and users to use the cloud storage systems on secure infrastructures and platforms that maintain their customers/own data privacy.

General Terms
Cloud systems, Security, Storage.

Keywords
Amazon web services, IaaS, PaaS.

1. INTRODUCTION
This paper, presents the main points of what providers and customers need to know about the cloud storage. From the definition of cloud computing systems, types, the cloud storage, the essential characteristics, secure infrastructure and platform to the different techniques of cloud storage data security which includes:

- Data-Confidentiality: makes sure that the data is shared between authorized users.
- Data-Integrity: verifies that the data is complete and authentic.
- Data-Availability: ensures that all the authentic users of the system have a direct access for the data they need at any time. This should help the providers/consumers of the cloud storage system services to make sure whether this system is secure or not when they create/use it.

Cloud computing is a general expression for anything delivers hosted services on the internet. The computing resources delivered into this shared network as a service, the origin of cloud expression from using the cloud shaped symbols as abstract way for the complicated infrastructures contained in system diagrams and flowcharts. In order to understand the cloud computing, the users should know the core technologies that it is based on and its essential characteristics (that will be explained in the following section). The National Institute of Standards and Technology (NIST) [1] captured well definitions for both of them.

Cloud computing is based on several technologies including:

- (SaaS) Software as a service: enables the user to access and use the provider's applications on the cloud infrastructure, considered as web applications. The interaction is through front-end portal (e.g., Google apps).
- (PaaS) platform as a service: enables the user to develop or create applications on the provider's infrastructure. It does not have control of cloud system infrastructure severs. It may have the control of the application and its configuration settings.
- (IaaS) Infrastructure as a service: enable system provider to manage the consumer access to the system (e.g., Amazon web services, Flexiscale and GoGrid).

The cloud can be classified into several types depending on users' needs. These needs vary according to the demands of security and the risk levels that differ from clout type to another.

In the first type, public cloud sells service for users on the internet, currently the Amazon web services is the largest public cloud provider. Private cloud is a proprietary network or data center that provides its services to a limited number of customers, which is the second type. The third type is known as virtual private cloud, which is a kind of private cloud, but the service provider uses public resources to build the private cloud. In the public cloud, user may lose control of many facilities which may be required to be more secure, since user may exchange the same data file with unfamiliar persons. In the private cloud, user has this privacy property just like the private network or company's network. All users are authentic users and all files on servers and storage devices under users' control. For the community cloud, it can be owned and managed by one or more community from the organizations that have shared concerns [1]. The hybrid cloud is a combination of one or more different type of cloud. The main goal for cloud computing is to provide scalable and easy access to computing resources and it services.

The cloud computing is being requested and used more and more especially in the IT industry, but the data is stored in public storage provider. That makes the users of cloud computing services care about more secure cloud storage systems. Due to the differences between enterprises, customers and individuals may need different levels of security and their ability to pay is different as well, so the cloud computing should present different levels of security according to the concept of computing on-demand service.

Large number of users and enterprises are using cloud storage systems. Cloud storage is simply used to store users' data and computations on a network of online storage. The data stored in virtual pools of storage that hosted by cloud storage providers. The most famous and trusted cloud storage providers now like Amazon cloud drive, I cloud, box, drop box and Google drive, let the users store their files for a free limited capacity or users can buy extra storage capacity. The fear of data corruption and systems damage or failure make the enterprises and individuals think about using the cloud storage systems, but they reluctant to do this. Due to the security concerns, users sacrifice the cloud storage privileges that they can gain.

This paper, discusses the main methodologies of securing the cloud storage. Also will present varied methods ranging from the secure platform as in [2] and [3] to the data security techniques comparison as in [4], [5], [6] and [7]. This study outlines should help the enterprises and individual users to verify their systems and data security.

The next section, mentions the essential characteristics for the cloud storage system. Knowing these characteristics will help the users to judge whether the system is secure or not, and at which level of security it is. The third section explains the securing methodologies of the trusted platform in [2] and [3]. The fourth section has an overview for three different techniques that reserve data confidentiality and integrity. Section 5 includes the comparison between the three techniques and how can the user choose the suitable option for the user's case. Section 6 will conclude the paper.

2. THE ESSENTIAL CHARACTERISTICS OF CLOUD STORAGE

This section, defines the essential characteristics of the cloud storage. According to NIST [1], the characteristics can be summarized the as follows:

- On-demand self-service: the consumers can do all their computing capabilities on their side without human interaction with the service provider
- Broad network access: cloud services can available to be accessed through the network using standard mechanisms.
- Resource pooling: the computing resources of cloud service pooled in multi-tenant model to be shared between all service users.
- Rabid elasticity: depending on users demands their capabilities scale rabidly up and down elastically.
- Measured services: optimize resource usage with making use of the metering capabilities.

The core of this paper is to reserve these essential characteristics with different levels of security (may be as high as possible) depends on the consumer demand.

To create or use cloud storage system, customers should ensure applying these essential characteristics. Considering the cloud characteristic Vulnerabilities [7] that should be avoided. These vulnerabilities are simply to break the essential security characteristics in the cloud system. For example, unauthorized access to the management interfaces: the on-demand service needs a management interface that is accessible by service users, any unauthorized access may cause system damage considered as cloud characteristic Vulnerability.

3. SECURE PLATFORMS FOR SECURE INFRASTRUCTURES

Amazon infrastructure has Amazon EC2 and Amazon S3 platforms. This section, compares different platforms with different level of security on different infrastructures according to the users' demands. In [2] and [3], there are two different platforms with different security aspects.

3.1 Terra

Terra is a traditional and trusted platform that has a flexible architecture that can be run side-by-side with normal applications on general computing platform. The usage of *Trusted Virtual Machine Monitor (TVMM)* enables Terra of partitioning the tamper-resistant hardware into multiple separated VMs, and deals with every VM as a box on the single general-purpose platform. Terra is applying both of open and closed box mechanisms. The users can choose either open or closed box upon the systems demands. Open box is applied for the public general-purpose platforms and the closed box for more privacy and security aspects. Terra also provides upfront attestation between a remote party and the user of the closed box to define what the closed box is contains cryptographically and consumers can attest to the system providers to verify the service security before using the system. Reserves the data privacy, even the system administrator cannot access the users' closed box. Terra can effectively secure systems that have virtual machines running on a single host.

3.2 Trusted Cloud Computing Platform (TCCP)

TCCP has a pre-attestation privilege that users can attest to the IaaS provider and make sure whether the service is secure or not before launch their virtual machines. TCCP guarantees the users' virtual machines confidentiality and integrity by using closed box execution environment on the infrastructure as a service (IaaS) providers' system (eg. Amazon EC2). Preventing the provider privileged administrator from inspecting or tampering with customers' VM contents.

To recover the platform attestation mechanism problem, the provider can divert the users' VM to a running node out of the platform. TCCP has to provide a remote attestation mechanism that guarantees the consistency of the platform security properties backend. The today's IaaS backends enables the closed box mechanism through using the TCCP without changing their architecture. By using Trusted Virtual Machine Monitor (TVMM) and trusted coordinator (TC) as a trusted computing base of TCCP, TVMM that hosts the consumers' VMs and prevent authentic user to inspect or modify them. According to the TC protocols the TVMM can protect its integrity over time. TC keeps track of the system customers (as nodes) through marking the trusted VM as trusted node. TC also can deal with adding a new node to the system, deleting or even temporary shutdown for any node.

Table 1 shows the comparison between Terra and TCCP through 10 points of comparisons. These 10 points present the main features of the compared systems.

Table 1 Terra and TCCP comparison

	Terra	**TCCP**
Trusted	Yes	Yes
Public/private	Both	Private
Data conf.&integrity	Within private	Yes
Hardware dependant	Yes	No
Closed / open box	Both	Closed
Reset within sys.boot	Yes	No
Security pre-attestation	Yes	Yes
Applicable to other IaaS	Yes(H/W dependent)	Yes(without changing its architecture)
Cryptographically secured	Yes	No
Applying TVMM, TC	TVMM	Both

4. DATA SECURITY

Data security means protecting the systems and customers' information from the external intruders and prevents the insiders from any unauthorized access.

Encryption, decryption and password techniques are common methods for data privacy protection especially for that users share via unsecured network. These methods become inefficient ways with the cloud storage systems. The reason that the IaaS service provider for the cloud system may have the encrypted data and the decryption key at the same time, which makes system administrators able to access the customers' private data. Generally in the cloud computing environment, storage services providers should have all the capabilities to ensure the security of the customers' data, but the system should prevent the authorized system administrators from accessing clients' data.

We will discuss other techniques, which make data storage using cloud storage systems more secure. The following three sub-sections explain three different methodologies, as follows:

- The first methodology is using separated encryption and decryption services [5].
- The second ensure the data security in cloud storage [4].
- Finally, the self protecting document for cloud storage security [6].

4.1 Separated Encryption And Decryption Service From The Storage Service

Instead of using the common data encryption and decryption techniques, which are not suitable for cloud computing as mentioned before, this method can use them separately from the storage service [5]. Data must not be stored in plaintext (first party). While the process of the encryption or decryption completed (second party) and delivered to an application like customer Relationship Management (CRM service). All the computations must be deleted. This model emphasizes the authorization for the storage and encryption/decryption for users' data must be done between two different service providers. One of them has the key without the customers' data and the other one has the encrypted data without the key.

4.2 Data Security Framework

This methodology provides a platform that ensures the data security using service-level-agreement (SLA) as a common standard that manages the interaction (type of service, quality of that service, terms of payment, and probability of data loss) between the consumers and the service providers. The technologies used to ensure the data security can be divided into three parts:

Storage protect, Transfer protect, and Authentication. These technologies applied in the cloud storage provider.

The first part is Secure Storage. To ensure the storage protection, the method divides the data into several small pieces and save it into different places. The service provider may have more several data centers that are far away from each other. According to the way that the data divided into pieces and where these pieces stored in the cloud storage system can be classified from low-to-high storage file level security as follows:

- Single-server level: store all the data pieces on the same server, but the data damage probability is too high since data piece corruption will cause to the whole file to be damaged.
- Cross-server level: store more than one copy of the data files to more than one backup servers, when any file corruption occurs the system automatically redirect the request to another backup server.
- Cross-cabinet level: each block divided into layered block that has the data and the same data backup and also has the server backup strategy that prevent the cabinet failure since the servers usually support data communication via internal network.
- Cross-data center level: this level of document security is to avoid the major disasters and accidents by using the backup layer strategy and store backup data copies into different data centers may be have thousands of miles between them. This approach has the highest level of security, but in the real world storing the data into several places (data centers) increases the data stealing risk probability.

Table 2 shows a comparison between the four possible file storage security levels.

Table 2 Levels of file storage security

	Single-server	Cross-server	Cross-cabinet	Cross-data center
Number of Backup security levels	No	1	2	2
Single data piece failure effect on the system	Yes	No	No	No
Server failure effect on the system	Yes	Yes	No	No

The second part is the Secure Transfer. In cloud storage systems, the data usually stored far away from the end user, the data transfer via a network, optimization in position techniques should be used to verify the data availability in any time by the end user. Either using a narrow subset of data computing needs for the programmable models or use the scheduling algorithms to list the system nodes as close to the storage system, or even use the content delivery network (CDN) to push the data near to the end users. In All approaches, the secure socket layer (SSL) is used to ensure the transfer process is safe. Cryptographic protocols such as Transport Layer Security (TLS) and secure socket layer (SSL) provide security for the communications through the networks (e.g. Internet). Encrypting the connection segments on the application layer by both (TLs) and (SSL) verify secure end-to-end transition at the transport layer.

The third and final part is authentication. In this method as mentioned in the transition mechanism the data should be encrypted, this will help in verifying the data privacy even from the system service administrator that considered having a privileged authentic access for all the system the administrator manages. The system provider has just an encrypted version of the user's data that cannot read, these data decrypted only by the data owner (end-user).

4.3 Self-Protecting Document

Providing secure information systems is difficult, but this difficulty being increasing when more than one organization work on one project and the data should be available and secure. The data confidentiality and integrity should be more difficult when the information about this project handed off to others outside these organizations.

This technique provides an approach that the document itself can reserve its privacy and security even when being exchanged on unsecured networks. Some security components like storage, access, and usage control – that the companies may deploy an information system to be responsible for – are encapsulated (encapsulation object-oriented concept) within the document to ensure autonomic document architecture for Enterprise Digital Right Management (E-DRM). This can't only be applied for files that can be exchanged through uncontrollable network like cloud computing systems, but also can be applied for the USB flash drivers.

Using the traditional authentic user name and password, there is a probability that the intruder can have access to these systems' files. The detailed works about cloud storage systems' security issues are introduced in [4], [5], [7] and in more details in [8].

A drawback may appear of using the centralized architectures for information sharing. As an example, the case is to connect the servers to ensure the security privacy but may that causes a bottleneck as it must check the status of all the users, and this force the user to install some applications on his side that shows its authentic access for the server. To overcome this drawback, the authors proposed to embed the data warehouse and the security modules in the same document to be self-protected. The server will be only responsible for the synchronization mission of the files being exchanged by the system users.

For the storage: using data base for storing the data contents as a node, each node has its information and the attached metadata (used to bind all the information of the customer's document together). This metadata provide permission to calculate different performance indicators to monitor the partner's operations on the document. The security control rules may prevent access parts in the document depending on its location.

For usage control: Using the security kernel enforces applying the security policy and monitoring all the operations that may perform on the document within the essential security modules (OrBAC for access, usage control, metadata management...etc).

The authentication: ensure authentic access can be performed by storing the license of the user outside the document, this license enforce applying the authentication security that prevent anyone but the authentic user to access the self-protected document. Others can access the file but with hidden parts that have higher security privilege dependent on its location in the document. Finally this approach is more useful for companies that need secure information systems for their documents that also needed to be available for the users and other information systems outside the company. Table 3 shows a comparison between the three cloud storage system's data security techniques.

5. CONCLUSION

This paper, starts from the very beginning introduction to the cloud computing and its types, cloud storage and its essential characteristics. Also explains two examples of trusted secure cloud storage platforms. Then, will show three different methodologies of cloud storage data security.

This paper provides the essential knowledge for users who want to use the cloud storage system, but may be hesitated because of security concerns. Into the bargain, the knowledge is useful for the users and providers to judge their storage system platform and data security especially at the time before they launch their systems' VM. The discussed storage security methodologies present comparisons between the latest researches that have been conducted in the cloud system platform security and data security.

Table 3 Comparison of three data security approaches for cloud storage

	Separated En/Dec from the cloud storage system	**Data security framework**	**Self-protecting document**
Trusted	Yes	Yes	Yes
Protection methodology	Enc/Decryption	Enc/Decryption	Document Encapsulation
Data confidentiality, integrity and traceability	Confidentiality and integrity	Confidentiality and integrity	Confidentiality, integrity and traceability
Authentication access technique	Authentication decryption key	Authentication access with automatic decryption by the system layers	Provided through the user license
Usage control for unauthentic users	Blocked	Untrusted users can't use the data	Limited and monitored by the security kernel
Usage recommendations	Business enterprises that have different storage providers(must be different)	Traditional cloud computing systems	The enterprises that has open source documentations available to inside and outside the system

6. REFERENCES

[1] Liu, F., Tong, J., Mao, J., Bohn, R., Messina, J., Badger, L., & Leaf, D. (2011). NIST Cloud Computing Reference Architecture. NIST Special Publication, 500:292.

[2] Garfinkel, T., Pfaff, B., Chow, J., Rosenblum, M., & Boneh, D. (2003, October). Terra: A virtual machine-based platform for trusted computing. In *ACM SIGOPS Operating Systems Review* (Vol. 37, No. 5, pp. 193-206). ACM.

[3] Santos, N., Gummadi, K. P., & Rodrigues, R. (2009, June). Towards trusted cloud computing. In *Proceedings of the 2009 conference on Hot topics in cloud computing* (pp. 3-3). USENIX Association.

[4] Zhang, X., Du, H. T., Chen, J. Q., Lin, Y., & Zeng, L. J. (2011, May). Ensure Data Security in Cloud Storage. In *Network Computing and Information Security (NCIS), 2011 International Conference on* (Vol. 1, pp. 284-287). IEEE.

[5] Hwang, J. J., Chuang, H. K., Hsu, Y. C., & Wu, C. H. (2011, April). A business model for cloud computing based on a separate encryption and decryption service. In *Information Science and Applications (ICISA), 2011 International Conference on* (pp. 1-7). IEEE.

[6] Munier, M., Lalanne, V., & Ricarde, M. (2012, June). Self-Protecting Documents for Cloud Storage Security. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on* (pp. 1231-1238). IEEE.

[7] Grobauer, B., Walloschek, T., & Stocker, E. (2011). Understanding cloud computing vulnerabilities. *Security & Privacy, IEEE*, *9*(2), 50-57.

[8] Dahbur, K., Mohammad, B., & Tarakji, A. B. (2011, April). A survey of risks, threats and vulnerabilities in cloud computing. In *Proceedings of the 2011 International Conference on Intelligent Semantic Web-Services and Applications* (p. 12). ACM.