# Securing Data Storage in Cloud Computing

Hyun-Suk Yu[1], Yvette E. Gelogo[2], Kyung Jung Kim[3]

## Abstract

Cloud computing is a new computing paradigm that attracted many computer users, business, and government agencies. Cloud computing brought a lot of advantages especially in ubiquitous services where everybody can access computer services through internet. With cloud computing, there is no need of physical hardware or servers that will support the company's computer system, internet services and networks. One of the core services provided by cloud computing is data storage. In the past decades, data storage has been recognized as one of the main concerns of information technology. The benefits of network-based applications have led to the transition from server-attached storage to distributed storage. Based on the fact that data security is the foundation of information security, a great quantity of efforts has been made in the area of distributed storage security. In this paper, the authors tried to study the threats and attacks that possibly launch in cloud computing data storage and proposed a security mechanism.

Keywords : Cloud Computing, Secret Key Sharing

## 1. Introduction

Cloud computing is a new computing paradigm where in computer processing is being performed through internet by a standard browser [1]. Cloud computing builds on established trends for driving the cost out of the delivery of services while increasing the speed and agility with which services are deployed. It shortens the time from sketching out application architecture to actual deployment. Cloud computing incorporates virtualization, on-demand deployment, Internet delivery of services, and open source software [2]. The Cloud Computing Architecture of a cloud solution is the structure of the system, which comprises on-premise and cloud resources, services, middleware, and software components, geo-location, the externally visible properties of those, and the relationships between them. The term also refers to documentation of a system's cloud computing architecture. Documenting facilitates communication between stakeholders, documents early decisions

Received(April 24, 2012), Review request(April 25, 2012), Review Result(1st: May 09, 2012, 2nd: May 23, 2012) Accepted(June 30, 2012)

[1] 565-701, Department of Early Childhood Special Education, Woosuk University, Ph.D.
email: dbgustnrsla@hanmail.net

[2] 306-791, Department of Multimedia Engineering, Hannam University
email: vette_mis@yahoo.com

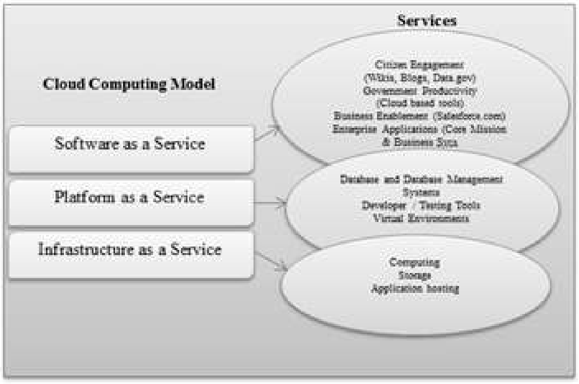
[3] (Corresponding author)565-701, Department of Child Welfare, Woosuk University, Professor.
e-mail: kkjung00@hanmail.net

about high-level design, and allows reuse of design components and patterns between projects [3]. The benefits of cloud computing are many. One is reduced cost, since you pay as you go. Other benefits are the portability of the application is that users can work from home, work, or at client locations. This increased mobility means employees can access information anywhere they are. There is also the ability of cloud computing to free-up IT workers who may have been occupied performing updates, installing patches, or providing application support. Along with the good services of Cloud Computing has to offer, there are security problems which make users anxious about the safety, reliability and efficiency of migrating to cloud computing. Big companies have second thought whether to move into the cloud because they might compromise the operation and the important information of the company. After analyzing and calculating the possible risk. Migrating into the "Cloud" will make computer processing much more convenient to the users. One of the considerations when moving to cloud is the security problems.

One consideration is that the unique issues associated with cloud computing security have not been recognized. Some researchers think that cloud computing security will not be much different from existing security practices and that the security aspects can be well-managed using existing techniques such as digital signatures, encryption, firewalls, and/or the isolation of virtual environments, and so on.

## 2. Background

### 2.1 Cloud Computing Models


Services
Cloud Computing Model
Software as a Service
Platform as a Service
Infrastructure as a Service
Citizen Engagement (Wikis, Blogs, Data.gov) Government Productivity (Cloud based tools) Business Enablement (Salesforce.com) Enterprise Applications (Core Mission & Business Sys)
Database and Database Management Systems Developer / Testing Tools Virtual Environments
Computing Storage Application hosting

[Fig 1] Cloud Computing Models

- ***a. SaaS:*** To use the provider's applications running on a cloud infrastructure and accessible from various client devices through a thin client interface such as a Web browser.
- ***b. PaaS:*** To deploy onto the cloud infrastructure consumer-created applications using programming languages and tools supported by the provider(java, python, .Net)

***c. IaaS:*** To provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications.

## 2.2 Layers of Cloud computing model

There are five layers in cloud computing model, the Client Layer, Application Layer, Platform layer, Infrastructure layer and server layer. In order to address the security problems, every level should have security implementation.

***Client Layer:*** In the cloud computing model, the cloud client consist of the computer hardware and the computer software that is totally based on the applications of the cloud services and basically designed in such way that it provides application delivery to the multiple servers at the same time, as some computers making use of the various devices which includes computers, phones, operating systems, browsers and other devices.

***Application layer:*** The Cloud application services deliver software as a service over the internet for eliminating the need to install and run the application on the customer own computers using the simplified maintenance and support for the people which will use the cloud interchangeably for the network based access and management of the network software by controlling the activities which is managed in the central locations by enabling customers to access the applications remotely with respect to Web and application software are also delivered to many model instances that includes the various standards that is price, partnership and management characteristics which provides the updates for the centralize features.

***Platform layer:*** In the cloud computing, the cloud platform services provides the common computing platform and the stack solution which is often referred as the cloud infrastructure and maintaining the cloud applications that deploys the applications without any cost and complexity of the buying and managing the hardware and software layers.

***Infrastructure layer:*** The Cloud Infrastructure services delivers the platform virtualization which shows only the desired features and hides the other ones using the environment in which servers, software or network equipment are fully outsourced as the utility computing which will based on the proper utilization of the resources by using the principle of reusability that includes the virtual private server offerings for the tier 3 data center and many tie 4 attributes which is finally assembled up to form the hundreds of the virtual machines.

***Server layer:*** The server layer also consist of the computation hardware and software support for the cloud service which is based on the multi-core processors and cloud specific operating systems and coined offerings.

## 2.3 Database Management in the Cloud

In recent years, database outsourcing has become an important component of cloud computing. Due to the rapid advancements in a network technology, the cost of transmitting a terabyte of data over long distances has decreased significantly in the past decade. In addition, the total cost of data management is five to ten times higher than the initial acquisition cost. As a result, there is a growing interest in outsourcing database management tasks to third parties that can provide these tasks for much lower cost due to the economy of scale. This new outsourcing model has the benefits of reducing the cost for running Database Management System (DBMS independently [1].

A Cloud database management system (CDBMS) is a distributed database that delivers computing as a service instead of a product. It is the sharing of resources, software, and information between multiply devices over a network which is mostly the internet. It is expected that this number will grow significantly in the future. An example of this is Software as a Service, or SaaS, which is an application that is delivered through the browser to customers. Cloud applications connect to a database that is being run on the cloud and have varying degrees of efficiency. Some are manually configured, some are preconfigured, and some are native. Native cloud databases are traditionally better equipped and more stable that those that are modified to adapt to the cloud.

Despite the benefits offered by cloud-based DBMS, many people still have apprehensions about them. This is most likely due to the various security issues that have yet to be dealt with. These security issues stem from the fact that cloud DBMS are hard to monitor since they often span across multiple hardware stacks and/or servers. Security becomes a serious issue with cloud DBMS when there's multiple Virtual Machines (which might be accessing databases via any number of applications) that might be able to access a database without being noticed or setting off any alerts. In this type of situation a malicious person could potentially access pertinent data or cause serious harm to the integral structure of a database, putting the entire system in jeopardy.

## 2.4 Cryptography

Cryptographic hash functions are an important tool of cryptography and play a fundamental role in efficient and secure information processing. A hash function processes an arbitrary finite length input message to a fixed length output referred to as the hash value. As a security requirement, a hash value should not serve as an image for two distinct input messages and it should be difficult to find the input message from a given hash value. Secure hash functions serve data integrity, non-repudiation and authenticity of the source in conjunction with the digital signature schemes. Keyed hash functions, also called message authentication codes (MACs) serve data integrity and data origin authentication in the secret key setting. The building blocks of hash functions can be designed using block ciphers, modular arithmetic or from scratch. The design principles of the popular Merkle–Damgard construction are followed in almost all widely used standard hash functions such asMD5 and SHA-1 [8].

## 3. Cloud Computing Attacks

As more companies move to cloud computing, look for hackers to follow. Some of the potential attack vectors criminals may attempt include:

a. ***Denial of Service (DoS) attacks:*** Some security professionals have argued that the cloud is more vulnerable to DoS attacks, because it is shared by many users, which makes DoS attacks much more damaging.

b. ***Side Channel attacks:*** An attacker could attempt to compromise the cloud by placing a malicious virtual machine in close proximity to a target cloud server and then launching a side channel attack.

c. ***Authentication attacks:*** Authentication is a weak point in hosted and virtual services and is frequently targeted. There are many different ways to authenticate users; for example, based on what a person knows, has, or is. The mechanisms used to secure the authentication process and the methods used are a frequent target of attackers.

d. ***Man-in-the-middle cryptographic attacks:*** This attack is carried out when an attacker places himself between two users. Anytime attackers can place themselves in the communication's path, there is the possibility that they can intercept and modify communications.

e. ***Inside-job:*** This kind of attack is when the person, employee or staffs who is knowledgeable of how the system runs, from client to server then he can implant malicious codes to destroy everything in the cloud system.

## 4. Security Requirements

Security measures assumed in the cloud must be made available to the customers to gain their trust. There is always a possibility that the cloud infrastructure is secured with respect to some requirements and the customers are looking for a different set of security. The important aspect is to see that the cloud provider meets the security requirements of the application and this can be achieved only through 100% transparency. Open Cloud Manifesto exerts stress on transparency in clouds, due the consumer's apprehensions to host their applications on a shared infrastructure, on which they do not have any control

In order to have a secured Cloud computing deployment, we must consider the following areas, the cloud computing architecture, Governance, portability and interoperability, traditional security, business continuity and disaster recovery, data center operations, incident response, notification and remediation, Application Security, Encryption and Key management, identity and access management [1]. One if the reason why users are very anxious of the safety of their data being saved in the cloud is that they don’t know who is managing it while in the server of the cloud computing service provider. Typical users who use the cloud computing service like storing their files on the server to access it anywhere they want through internet, don’t bother much about the security of their files, those documents are common files that don’t need to be secured. But in the case of big companies which have very important information to take care of, they need to have secured cloud computing system.

In order to have secure cloud system, the following aspect must be considered:

***Authentication:***

Authentication is the process of verifying a user or other entity’s identity. This is typically done to permit someone or something to perform a task. There is variety of authentication system, some are stronger than others. A strong authentication system ensures that the authenticators and messages of the actual authentication protocol are not exchanged in a manner that makes them vulnerable to being hijacked by an intermediate malicious node or person. That is, the information used to generate a proof of identity should not be exposed to anyone other than the person or machine it is intended for.

***Authorization:***

Authorization is when the system decides whether or not a certain entity be allowed to perform a requested task. This decision is made after authenticating the identity in question. When considering an authentication system for a particular application, it is crucial to understand the type of identifier required to provide a certain level of authorization.

***Confidentiality:***

Confidentiality is needed when the message sent contains sensitive material that should not be read by others and therefore must not be sent in a comprehensible format. A loss of confidentiality is the unauthorized disclosure of information. Confidentiality, as it relates to security and encryption techniques can be obtained by encrypting messages such that only intended recipient are able to read them.

***Integrity:***

Integrity is ensuring that the data presented are true and valid master source of the data and includes guarding against improper information modification or destruction to ensure information non-repudiation and authenticity. A loss of integrity is the unauthorized modification, insertion, or destruction of information.

One way of ensuring of data integrity is by using simple checksums which prevent an attacker from forging or replaying messages. Checksum is usually implemented when the channel between communication parties is not secure and ensure that the data has reached its destination with all bits intact, if bits have been modified, that the modification will not go unobserved.

***Non-Repudiation:***

Non-repudiation is ensuring that a traceable legal record is kept and has not been changed by a malicious entity. A loss on non-repudiation would result in the questioning of the transaction that has occurred. A simple example of non-repudiation is signing o contract. The signer cannot claim they did not agree a contract because there is an evidence that they did agree. The difference is that a signature can be forger but good encryption cannot.

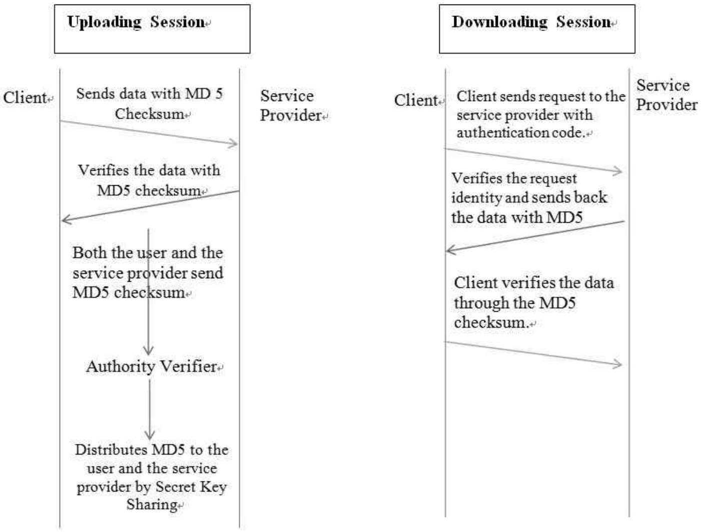
## 5. Proposed Security Mechanisms

The procedure is secure for each individual session. The integrity of the data during the transmission can be guaranteed by the SSL protocol applied. However, from the perspective of cloud storage services, data integrity depends on the security of operations while in storage in addition to the security of the uploading and downloading sessions. The uploading session can only ensure that the data received by the cloud storage is the data that the user uploaded; the downloading session can guarantee the data that the user retrieved is the data cloud storage recorded. Unfortunately, this procedure applied on cloud storage services cannot guarantee data integrity. To illustrate this, let's consider the following

In Uploading Session, user sends data to the service provider with MD5 checksum then the service provider verifies the data with MD5 checksum. Both the user and the service provider send MD5 checksum to Authority Verifier. Authority Verifier verifies the two MD5 checksum values. If they match, the Authority verifier distributes MD5 to the user and the service provider by Secret Key sharing. Both sides agree on the integrity of the uploaded data and share the same MD5 checksum by secret key sharing, and the Authority verifier owns their agreed MD5 signatures.

In Downloading Session, client sends request to the service provider with authentication code. Then Service Provider verifies the request identity, if it is valid, the service provider sends back the data with MD5 checksum. Client verifies the data through the MD5 checksum.

When disputation happens, the user or the service provider can prove their innocence by checking the shared MD5 checksum together. If the disputation cannot be resolved, they can seek further help from the Authority verifier for the MD5 checksum. Here are the special cases. When the service provider is trustworthy, only the user needs the MD5 checksum; when the client is trustworthy, only the service provider needs MD5 checksum; if both of them can be trusted, the Authority verifier is not needed. This is the method used in the current cloud computing platform.


Uploading Session
Client
Sends data with MD 5 Checksum
Service Provider
Verifies the data with MD5 checksum
Both the user and the service provider send MD5 checksum
Authority Verifier
Distributes MD5 to the user and the service provider by Secret Key Sharing
Downloading Session
Client
Client sends request to the service provider with authentication code.
Service Provider
Verifies the request identity and sends back the data with MD5
Client verifies the data through the MD5 checksum.

[Fig 2] Proposed mechanism

## 6. Conclusion

One of the core services provided by cloud computing is data storage. This poses new challenges in creating secure and reliable data storage and access facilities over remote service providers in the cloud. The security of data storage is one of the necessary tasks to be addressed before the blueprint for cloud computing is accepted. In this paper we discussed the security requirements of cloud computing data storage security and the solutions for the security problems.

References

[1] Masayuki Okuhara et al, “Security Architecture for Cloud Computing”, FUJITSU Sci. Tech. J., Vol. 46, No. 4, pp. 397-402 (October 2010)

[2] Sun Microsystems, Inc., “Introduction to Cloud Computing Architecture”, White Paper, 1st Edition, June 2009

[3] Gerald Kaefer, “Cloud Computing Architecture”, Corporate Research and Technologies , Munich, Germany, Siemens AG 2010, Corporate Technology

[4] Peter Tseronis, “Cloud Computing Overview: A Federal Government and Agency Perspective”, ArchitecturePlus Seminar -Cloud Computing, Web 2.0 and Beyond: A Vision of Future Government Operations, August 13, 2009

[5] Kangchan Lee, “Cloud Computing”, Vice Chairman of ITU-T FG Cloud Chairman of Mobile Cloud WG in CCF in Korea, ETRI.

[6] VeriSign, “Digital ID, A Brief Overview”, A VeriSign White Paper, 2004 VeriSign,

http://www.verisign.com/static/005326.pdf

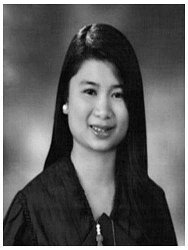
[7] Rajkuma Buyya, James Broberg, Andrzej Goscinski, "Cloud Computing Principles and Paradigms" A John Wiley & Sons, Inc. Publication, 2011

[8] Peter Stavroulakis, Mark Stamp, “Handbook of Information and Communication Security”, Springer Heidelberg Dordrecht London NewYork, ISBN 978-3-642-04116-7.

## Authors

**Hyun-Suk Yu**

1990 : B. S. in Child Welfare, Woosuk University
2006 : M. S. in Child Welfare, Woosuk University
2012 : Ph.D. in Early Childhood Special Education, Woosuk University
Research Interests: Special Education Technology



**Yvette Gelogo**

2006~2010 Bachelor of Science in Information Technology, Western Visayas College of Science and Technology.
Currently, Masters for Multimedia Engineering, Hannam University.
Research Interests: Ubiquitous Healthcare System Development, Mobile System Development and Design, Information Security, SCADA Security, Ubiquitous Learning, Biometric Authentication

**Kyung Jung Kim**

1980 : B. S. in Education of Teacher, Chungang University.
1983 : M. S. in Early Childhood Education, Chungang University.
1989 : Ph.D.in Early Childhood Education, Chungang University
1985 : Research Professor: Internationale Jugend Bibliotek(IJB), Institution in Germany
1999 : Visiting Scholar: Michigan State University in USA.
Currently : Professor, Department of Child Development & Welfare, Woosuk University, and President of Association of Korea Family Educare (from 1994)
Research interests : Child welfare, Child care information protection.