A survey on Data Confidentiality in Cloud Storage environments

Nyengele Nkota André
M.Tech Scholar, Department of Computer Science and Engineering
School of Computing and Information Technology
Reva University
Bangalore, India
sannynkota@gmail.com

Bijay Kumar Jyotishi
Associate Professor, Department of Computer Science and Engineering
School of Computing and Information Technology
Reva University
Bangalore, India
bijayj@revainstitution.org

Abstract-Cloud storage service, an Internet hosting service aimed at offering consumers the possibility to upload their files to the service providers' facilities and access them at any time, from anywhere via Internet, becomes a better way of providing individuals and organizations with a set of services and facilities to store and/or share their files. As such, it has led to a specialised cloud computing business and service delivery model known as Storage-as-a-Service. However, since cloud consumers' important files have to be stored in the cloud storage service providers' premises, the confidentiality of these stored files becomes a real issue. In this paper, different solutions to solve this issue have been proposed.

Keywords: ***Cloud storage, cloud storage security, confidentiality, data encryption.***

I INTRODUCTION

Cloud Computing is considered nowadays as a far better way of providing individuals and organizations with a set of both organized and structured services needed by the latter with different benefits such as cost optimization, omnipresent access of their information, and many more. Apart from the basic and fundamental Cloud Delivery Models (CDMs) used in Cloud Computing technology, i.e. Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS), different other specialized cloud service delivery models do exist based on the aforementioned. They may be: Storage-as-a-Service, Database-as-a-Service, Security-as-a-Service, Communication-as-a-Service, Integration-as-a-Service, Testing-as-a-Service, and Process-as-a-Service. [1]

Though all these services are available and many more are expected to be created in near future, many potential cloud consumers are still not totally relying on these ones because of different security-related issues they present. These issues may be related to data privacy (as storing organization's business data in a cloud provider's IT resources requires a high level of confidence in this one), some other issues are related to data storage, and some may even be related to data ownership.

The Storage-as-a-Service (STaaS will be used instead of SaaS to avoid confusion with Software-as-a-Service in this paper), from the definition given in [2] and [3], and the description made in [1], can be understood as both a *business* and a *service delivery model* in which a Storage Service Provider (SSP) leases or rents its storage infrastructure to a Storage Service Consumer (SSC) for the latter to store its data. That means, the SSP is a company owning IT resources that can be used by other companies or individuals (SSC) for storing their private and/or business data.

The service, as for other cloud service delivery models, is meant to meet a number of basic cloud services requirements such as: ubiquitous access, on-demand usage, resiliency, elasticity, etc. These requirements are guaranteed at the time both SSP and SSC hold a service contract known as Service Level Agreement (SLA). It (the service) can be consumed either by subscribing to cost-per-gigabyte-stored and/or cost-per-data-transferred basis, or just as a free storage (generally in exchange of advertising services in a way or in another).

With reference to STaaS business/delivery model, data can be stored and managed in different formats, i.e. objects, blocks, files, etc. When data is stored as objects, each object in the Object storage (or Object-based storage) typically includes the data itself, a variable amount of metadata, and a globally unique identifier. On the other side, when data is stored and managed as blocks, storage handles it as blocks within sectors and tracks. Additionally, with block storage, files are split into evenly sized blocks of data, each with its own address but with no additional information (or metadata) to provide more context for what that block of data is. [4]

Finally, when data is stored and managed as files, the storage uses file systems to handle these files as a file hierarchy. [5]

SNIA [6] has defined a reference model known as Cloud Data Management Interface to be used for both the basic functions of persistent storage i.e. *create*, *read*, *update*, and *delete* (CRUD) data in the cloud storage environment and such advanced functional features as:

- Allow clients to discover the capabilities available by the cloud provider;
- Manage containers and the data that is placed in them and;
- Allow metadata to be associated with containers and the objects they contain. [6]

According to this Reference Model (Figure 1), it can be noticed that for most types of data storage (block storage, object storage, file system, etc.) there is a container that is meant to hold the stored data and make referencing and retrieval operations easier as well as help in managing stored data in a more efficient way.

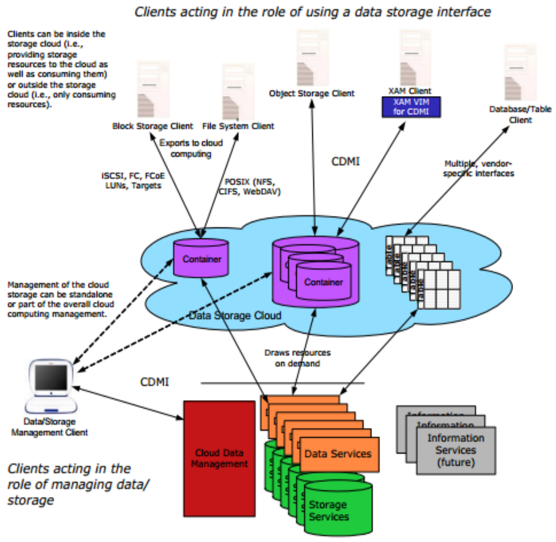
Clients acting in the role of using a data storage interface
Clients can be inside the storage cloud (i.e., providing storage resources to the cloud as well as consuming them) or outside the storage cloud (i.e., only consuming resources).
Block Storage Client
File System Client
Object Storage Client
XAM Client
XAM VIM for CDMI
Database/Table Client
Exports to cloud computing
iSCSI, FC, FCoE LUNs, Targets
POSIX (NFS, CIFS, WebDAV)
CDMI
Multiple, vendor-specific interfaces
Container
Container
Table
Data Storage Cloud
Management of the cloud storage can be standalone or part of the overall cloud computing management.
Draws resources on demand
CDMI
Data/Storage Management Client
Cloud Data Management
Data Services
Storage Services
Information Services (future)
Clients acting in the role of managing data/ storage

Figure 1: Cloud Storage Reference Model

Apart from this introductory and the conclusive sections, this paper work is divided into three main sections. The section II has discussed the different security issues encountered while subscribing to Storage-as-a-Service business and service delivery model to store one's data in the SSP's facilities. In the third section diverse solutions proposed for coping with the issues presented in the section II have been surveyed. But the main focus in this section has been put on the confidentiality because of its being the major security preoccupation for most organisations.

II SECURITY ISSUES IN CLOUD STORAGE

As mentioned above, though Cloud computing and all its related services offer a tremendous set of benefits to the cloud service consumers, there are still more to say about the security issues the present also. In this section, some of the well-known security issues related to cloud computing in general, and in particular to cloud storage, will be discussed.

1. Privacy and confidentiality

C. Kauba et al. [7] list a number of privacy and confidentiality related issues that can be encountered while using cloud computing and different services it offers, particularly cloud data storage. They group attackers in three major categories: hackers, cloud providers, and governments. Each of these three groups represents eventual privacy and confidentiality issues. For example, hackers generally have interest in cloud consumers' private data for illegal purposes. According to authors, credit card information, bank account details, health records, bank login details and so on, are some of the focus point of interest as hackers, once in possession of these details, will make illegal use of them by selling them or even by stealing owner's money in a bank account.

Cloud service providers are listed here as attackers by authors not to mean that their access to cloud consumers' private data is illegal, but rather that they leverage their legal access to this data for *achieving profit*. According to authors, "Many service providers scan user data on tags which are then used to show highly personalized ads." Or again "more complex data and statistics are recorded, bundled, and analysed ([using] data mining) to be able to do so called user profile marketing, making prediction on what items user might buy in the near future, what is next travel destination, etc."

The third group, governments, is mentioned here as they have, according to authors, *the most extensive legal possibilities to access private user data stored in the cloud.* In most cases, the objective is to combat the crimes and terrorist activities. But law enforcement can be made to cause cloud service providers to disclose cloud consumers' protected or encrypted data to them. They can even ask the data encryption keys to be delivered to them for further actions whenever required.

2. *Availability and Accessibility*

The availability of an IT resource (which is usually expressed as a percentage of *up-time* and measured as *Total Up-time/ Total time*) [1] is a supposed to be of 100% (which is not always the case in the cloud computing environment). Since data stored in cloud storage has to be accessed via Internet (in most cases), the performance-related issues will likely be experienced. Also, without accessing the WAN (Wide Area Network), it is impossible to access and use one's data. But most important is the capacity of the SSP to guarantee all required mechanisms to provide a highly available service, i.e. a service with a downtime exceptionally low. This can be achieved for example by offering a multiplicity of data storage.

3. *Access control/ Data ownership*

As well known, when the cloud service consumer uploads his data to the cloud service provider's premises (here we address the case where the SSP is other than the user's organisation for example), he/she loses control over his or her data, because it is stored on a computer belonging to someone [15]. This may result in conflicts as the privacy may be of bigger concern for the actual data owner, but not the case for the SSP who may want to make business profit over this very data. [16] Additionally, because data is stored in someone's facilities, it is obviously impossible to control who (Government, SSP, …) will eventually access it, when and for what purpose.

III LITERATURE SURVEY

With regard to the above-mentioned issues, particularly those related to privacy and confidentiality of data stored in cloud storage environments, a great number of solutions (most of them implying cryptographic techniques) have been proposed. In this section, a review of some of these solutions has been done.

1. *Cloud Storage as the Infrastructure of Cloud Computing [8]*

In this article the authors gave a quick introduction to cloud storage. They covered the key technologies in Cloud Computing and Cloud Storage, several different types of clouds services, and even described the advantages and challenges of Cloud Storage after the introduction of the Cloud Storage reference model.

2. *Applying Encryption Algorithm for Data Security in Cloud Storage [9]*

The authors proposed a simple, secure, and privacy-preserving architecture for inter-Cloud data sharing based on an encryption/decryption algorithm aiming at protecting the data stored in the cloud from the unauthorized access. The proposed solution was a two-phase solution. It made use of Advanced Encryption Standard (AES) algorithm for encryption/decryption task. After encrypting data using AES key before uploading it to cloud, the latter was encrypted using RSA algorithm so as to provide some control in data access, i.e., in case an unauthorized user tries to access data, he first needs to have the RSA key to decrypt the AES key and then decrypt data itself.

3. *An Enhanced Security Technique for Storage of Multimedia Content over Cloud Server [10].*

In this paper, authors mainly proposed framework to provide data storage in the cloud environment with secure user cloud security. They presented a three tier architecture in which original file (text, audio, video, image) was stored on local server, the encrypted (with RSA and Twofish) filename and description of the original file were stored on cloud server, and the encryption/decryption key was stored in Gmail account.

4. *Secure and efficient Cloud computing framework [11]*

In this paper, authors proposed an efficient and secure framework for cloud computing storage environment. The framework they proposed established distinction between different users' data based on their confidentiality level. They categorized the data into three types depending on their importance, and based on that, they selected the appropriate encryption algorithm with the suitable key size to provide the required security level. This way, they reduced the cost and the complexity and shortened the time needed to store the data securely. Besides, they made use of AES128 and AES256 for the two levels (high and higher) of confidentiality from users' chosen classification.

5. *CloudSafe: Storing Your Digital Asset in the Cloudbased Safe[12]*

Authors proposed CloudSafe to enhance the availability and confidentiality of the stored information in the cloud through encrypting and encoding data into several cloud storage providers. In order to make a safe, dependable and quick data access repository possible, CloudSafe solution was intended to offer a cloud-based personal digital asset safe service which delivers the valuable assets between several cloud providers by using erasure coding and cryptography. According to authors, the availability improves thanks to using erasure coding to distribute the data on several cloud providers, in order to recover data access when a provider fails. AES algorithm has been used for encrypting and decrypting data to keep data confidentiality.

6. *Implementing Digital Signature with RSA Encryption Algorithm to Enhance the Data Security of Cloud in Cloud Computing [13]*

In this paper, authors tried to evaluate cloud storage methodology and data's security level in the cloud through implementing digital signature with RSA algorithm to encrypt the data while it is being transferred. They defined digital signature as an arithmetical scheme for proving digital message authenticity. Implementing a valid digital signature with RSA algorithm guarantees data security in the cloud and allows the recipient to verify that the message was made by a recognized sender.

7. *Efficient and secure data storage in Cloud Computing through Blowfish, RSA, and Hash Function[14]*

Authors presented secure file exchanging on Cloud using Blowfish, RSA, and Hash algorithms for solving data security (confidentiality), authentication, and integrity problems of files on the cloud. Data security was improved by using cryptography algorithms. The rightness of data was verified by introducing Hash techniques. Enhanced system (Blowfish + RSA + Hash value) compared with simple RSA and Blowfish on basis of some performance parameters like throughput, encryption time, ciphertext, and delay time.

The *Table 1* gives a comparative survey of all the solutions reviewed in this section. It includes the title of each paper, the year it was published, the security issues addressed in it, cryptographic algorithms used for encryption/decryption (wherever applicable), along with different advantages and disadvantages of each of them.

TABLE I. COMPARATIVE SURVEY ON DATA CONFIDETIALITY IN CLOUD STORAGE ENVIRONMENTS

No	**Title**	**Publication year**	**Addressed security aspects**	**Used cryptographic algorithms**	**Advantages and disadvantages of the proposed methods**
1	Cloud storage as the infrastructure of Cloud computing	2010	Issues in Cloud storage	-	**Advantage** Clear understanding of what cloud storage is, differed benefit coming from the use of all implied technologies
2	Implementing Digital Signature with RSA Encryption Algorithm to enhance the data security of Cloud in Cloud computing	2010	Access control and authentication	RSA	**Advantage** Secure access to one's stored data in cloud storage via authentication. **Disadvantage** No confidentiality issue addressed.
3	CloudSafe: Storing Your Digital Asset in the Cloud-based Safe	2013	Data confidentiality Availability	AES	**Advantage** Use of AES for encryption **Disadvantage** The whole data is encrypted with same key size, thus resulting in performance issues
4	An Enhanced Security Technique for Storage of Multimedia Content over Cloud Server.	2013	Multimedia files confidentiality, Authentication	RSA Twofish	**Advantage** No performance issue as data itself is not encrypted **Disadvantage** Makes privacy easy to compromise.
5	Secure and efficient Cloud computing framework	2015	Data confidentiality Data integrity	AES128 AES256 SHA-1,…	Advantage Optimization of encryption/decryption related performance as data is encrypted based on its level of confidentiality

6	Efficient and secure data storage in Cloud Computing through Blowfish, RSA, and Hash Function	2015	Data integrity Data confidentiality Authentication	Blowfish RSA Hash algorithm	**Advantage** Data confidentiality and data integrity addressed. **Disadvantage** Whole data has to be encrypted, thus comes performance issue
7	Applying Encryption Algorithm for Data Security in Cloud Storage	2016	Data Confidentiality Access control	RSA AES	**Advantage** Data confidentiality and access control provided **Disadvantage** Performance issue for encryption whole data

IV. CONCLUSION

Cloud storage is a very important aspect of cloud computing and its services. It ships with both benefits and disadvantages. In this paper, some of the recent issues encountered in cloud storage environments have been reviewed. Also, some of the countermeasures applying to confidentiality and privacy of data stored in cloud storage as proposed by different authors have been surveyed. A brief comparative analysis has been done to enlighten advantages and disadvantages of the proposed solutions.

As part of the existing challenges, some issues still are not addressed in the proposed solutions. For example, most do not consider the fact that, for sure confidentiality, *only* the data owner and eventual authorised data users should be in possession of the data encryption keys. Thus encryption should be done from the data owner's system. Also, for those considering this solution, there either a lack of performance in the encryption/decryption process as even non-confidential data is also encrypted (which is not required at all). The future scope will be to find a solution where performance-related issues are taken into consideration while providing a safer way to keep stored data confidential.

REFERENCES

[1] Thomas Erl, Ricardo Puttini, Zaigham Mahmood, ***Cloud Computing: Concepts, Technology and Architecture***, PHI, 2013, pp.63-68,406.

[2] Techopedia. **Storage as a Service (SaaS)**. Retrieved from https://www.techopedia.com/definition/24900/storage-as-a-service-saas. 15/04/2017.

[3] Techtarget. **Storage as a Service (SaaS)**. Retrieved from http://searchstorage.techtarget.com/definition/Storage-as-a-Service-SaaS. 15/04/2017.

[4] Porter De Leon, Yadin; Tony Piscopo. "Object Storage versus Block Storage: Understanding the Technology Differences". Retrieved from https://www.druva.com/blog/object-storage-versus-block-storage-understanding-technology-differences/. *15/04/2017.*

[5] Wikipedia. **Object Storage**. Retrieved from https://en.wikipedia.org/wiki/Object_storage. 15/04/2017

[6] Storage Networking Industry Association. "Cloud data management interface (CDMI)." (August 2014)

[7] C. Kauba, "When the clouds disperse. Data Confidentiality and Privacy in Cloud Computing" in SE seminar of Informatics, Dept. of Comput. Sci., University of Salzburg, Salzburg, Austria, Jul. 14, 2013, pp.12-18.

[8] J. Wu, L. Ping, X. Ge, Y. Wang, and J. Fu, "Cloud storage as the infrastructure of cloud computing," in *Intelligent Computing and Cognitive Informatics (ICICCI), 2010 International Conference on.* IEEE, 2010, pp. 380–383.

[9] Kartit, Z., Azougaghe, A., Idrissi, H. K., El Marraki, M., Hedabou, M., Belkasmi, M., & Kartit, A. (2016). Applying Encryption Algorithm for Data Security in Cloud Storage. In *Advances in Ubiquitous Networking* (pp. 141-154). Springer Singapore.

[10] Gupta, P. and Brar, A.K., 2013. An Enhanced Security Technique for Storage of Multimedia Content over Cloud Server. *International Journal of Engineering Research and Applications (IJERA), 3*(4), pp.2273-2277.

[11] Lo'ai, A. T., Al-Qassas, R., Darwazeh, N., Jararweh, Y., & AlDosari, F. Secure and Efficient Cloud Computing Framework. In *the Proceedings of the IEEE International Conference on Cloud and Autonomic computing (Cloud Cyber Security workshop)* (pp. 291-295).

[12] Zhang, Q., Luo, B., Shi, W., & Almoharib, A. M. (2013). Cloudsafe: Storing your digital asset in the cloud-based safe. *Wayne State University, Detroit, USA, Tech. Rep.*

[13] U. Somani, K. Lakhani, and M. Mundra, "Implementing digital signature with RSA encryption algorithm to enhance the data security of cloud in cloud computing," in *Parallel Distributed and Grid Computing (PDGC), 2010 1st International Conference on*. IEEE, 2010, pp. 211–216.

[14] Kaur, N., & Singh, H. ***Efficient and Secure Data Storage in Cloud Computing Through Blowfish, RSA and Hash Function***, International Journal of Science and Research (IJSR), 4(5), May 2015,

[15] Haghighat, M., Zonouz, S., & Abdel-Mottaleb, M. (2015). **CloudID: Trustworthy Cloud-based and Cross-Enterprise Biometric Identification.** Expert Systems with Applications, 42(21), 7905–7916.

[16] Wikipedia. **Cloud computing issues**. Retrieved from https://en.wikipedia.org/wiki/Cloud_computing_issues. 24/04/2017.