International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 2, February- 2013

Data-a-Service and their Security Concerns in Cloud

Vivek Bhushan
Amity University
Noida, India

Aarti Khetan
Amity University
Noida, India

Subhash Chand Gupta
Amity University
Noida, India

***Abstract*-**In the field of Information technology (IT), the next generation is called as fourth generation technology. With the advancement in IT there has been an exponential growth in the amount of data produced in every field like educational organization, computational results, material science, social science, cosmology, astronomy, medicine and climate. Cloud computing provides a solution for handling such a vast amount of data by providing an unlimited space where a user can store his/her data and access that data from anywhere using Internet. But in spite of the cloud advantages it also suffers from several vulnerabilities like data integrity, confidentiality and maintainability. These factors create fear in the mind of many business tycoons to move their business on the cloud platform. This paper brings a brief idea on the cloud data generation, several policies to maintain data security and the factors hindering the cloud adoption.

Keywords- cloud computing; data life cycle; data security; data protection as a service; security standards.

I. INTRODUCTION

Cloud computing has revolutionized the way of field of IT by accelerating services in the form of *Software-as-a-Service* (SaaS), *Platform-as-a-Service* (PaaS), and *Infrastrcture-as-a-Service* (IaaS). Due to these features it has attracted a large number of businesses ranging from high level to small scale industries. Cloud helps to reduce the infrastructure cost setup which helps many small scale industries to boom in the market.

With the advancement in IT, the amount of data is also increasing exponentially and has become *too big*. "Too big" refers to the increase in information produced by increasing organizations every year.

The most inhibiting factor for cloud adoption is the *data management*. Data management deals with the *confidentiality*, *integrity*, and *availability* of data in the cloud. Data confidentiality means that only legitimate user can use that information data. Data integrity means that the data which is kept in the cloud must not be modified by any other used except the owner of the file. Data availability means that the data must be present every time so as to retrieve quickly by the user without any delay.

To solve this problem many IT professionals have put their security model to protect data such as described as *Proof of Retrievability* [1]-[4], *Proof of Possession* [5]-[8], and *Data Signature* [9] to maintain integrity of data in the cloud. There are also works on *Data Auditing* [9], *Data Backup Mechanism* [10], and *Security Models* [12]-[13].

The solutions provided by most of the technology only deals with the one stage of data life cycle i.e., storage while data has many stages in its life cycle according to the data movement and various operations performed on it. Whether the data is kept on a cloud server but the data is transmitted over a network which is also a problem because of many untrusted networks which violates the data security. There is no any defined and fixed security policies related to the cloud because every countries has its own protocols and this is the core reasons for hindering cloud adoption.

II. LIFE CYCLE OF DATA

Every data has its own life cycle ranging from its generation to its destruction. Many of the professionals have defined the data life cycle based on data usage. Oracle [14] describes the data life cycle into three states - *active*, *less active*, and *archived* state according to the frequency of data access. After the creation of data it, is in active state for some time, as the time goes on it becomes less active and lastly it goes to archived state. Cloud security alliance [15] defined the data life cycle into six states – *produce*, *store*, *use*, *share*, *archive*, and *destruct.*

Here we define the data life cycle into seven stages – *generation*, *transfer*, *use*, *share*, *storage*, *archive*, and *destruct* as shown in figure 1.1

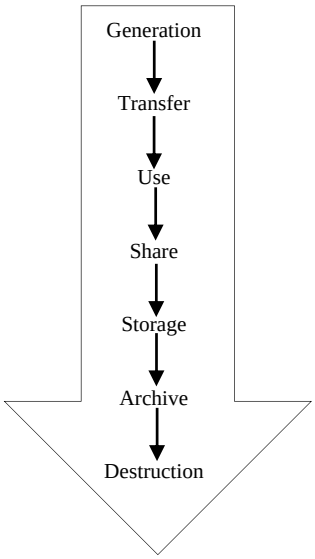

Figure 1.1 Life Cycle of Data in Cloud

The various phases of life cycle are described as follows-

1. *Data Generation –*

It is the process of data creation from scratch and deals with the data ownership. Data may be created either by the user or cloud server. If the data is to be moved into cloud, it must be kept in mind that who is going to own that data.

2. *Data Transfer –*

Within an organization, data transmission requires a little or no encryption. But when the data is to be moved outside an organization, the data management must be kept in mind like data confidentiality and data integrity so as to secure the data from being leaked and modified.

3. *Data Use –*

It refers to the way of extracting and using various types of data in cloud. The feasibility of encrypting a data depends on the type of data. For example, encryption of static data on Amazon S3 storage service is feasible. But in the case of cloud based applications using PaaS model, it is not feasible because it will create the problem of indexing and querying. In cloud computing the platform is shared which is accessed by many users. It creates the problem of multi-tenancy because the data of different users may be stored on the same virtual machine.

4. *Data Share –*

It is the way of increasing the data use range by providing permissions to other users to use that data but it leads to complexity. The owner "A" of the data "D" can authorize other user "B" to use the data "D" and in turn user "B" can further authorize user "C" to use data "D", thus expanding the range of data use. In order to share the data the owner of the data must agree upon the correct *service level agreements* (SLAs) in order to maintain their data.

5. *Data Storage –*

This is the way of storing the data in the cloud. A data in the cloud is stored on the basis of the platforms. The storing of data relates with three security concerns – confidentiality, integrity, and availability. The solution to this problem is strong encryption method and a very strong key management concept. But doing this, other factors must also be considered like efficiency of data encryption and processing speed.

6. *Data Archival –*

It deals with storage media for data archiving for a longer time. There is a risk of data leakage if it is stored on portable media. So, a user must ask their service provider about the media storage service and must mention in the SLAs. If this is not provided by the service provider the problem of data availability will be threatened.

7. *Data Destruction –*

It is completely erasing a data when it is no longer to be used. There are several data recovery techniques, so it must be taken in care that the data destroyed must not be recovered again. If it not taken care, the problem of disclosure of sensitive data can be raised.

III. DATA SECURITY

According to a survey conducted by The European Network and Information Security Agency (ENISA) on more than 70% of small medium enterprises [16], table 1 shows the importance of data security concerns.

Table 1. Data Security Importance

Condition	Importance (%)
Corporate data confidentiality	95
Privacy	88
Services/data availability	88
Services/data integrity	87
Control of services/data loss	75
Responsibility of service providers	72
Repudiation	57

A. Type of Data in Cloud –

There are basically three types of data [17] in the cloud system which is to be secured-

1. *Transmission Data-*

It is the data referring to the data being transmitted in a processing.

2. *Storage Data-*

It refers to the data which are either old or in archived state.

3. *Processing Data-*

It refers to the data on which certain operations are being performed.

B. Threats to Data Security -

There are two types of threat to data in the cloud – *internal* and *external* threats [18].

Internal Threat means that the threat is inside the cloud either by the cloud service provider or by the users in the cloud. As it is clear that the data of different users are kept on the same physical machine, so, the cloud service provider have full authority of access on the data of all the users of cloud. A cloud service provider can easily perform unauthorized operations on the user's data in the cloud as a legitimate user like modifying user's data, taking backups of the user's data, or move the user's data from one location to another without the user's permission.

External Threats means the threats arising outside cloud. It can be in the form of attacks like virus attack, man in middle attack, masquerading attack. All these attacks are done in order to steal the user's name and password so that any operations on data can be performed.

C. Data Security Challenges –

The challenges to data security [19] in cloud can be classified into following categories –

1. *Data Transmission-*

Data transmission is the biggest challenge to data in the cloud. It is very difficult to detect the stealing of transmission process when an enterprise is performing some transmission in the cloud. It may be the case when the data gets leaked and damaged or modified by a hacker.

2. *Data Migration –*

It is one of the hot topic in cloud computing from security point of view. As it is known that there is multi-tenancy in the cloud, so the problem of sensitive information still holds like military information if it is deployed in the cloud. The migration of data must be taken care carefully so that no any node in the cloud gets leaked.

3. *Data Remanence –*

It is the process of erasing the data from the cloud completely. When a user wishes to remove their data from cloud, the user must be able to verify that all the data which were stored on the cloud are completely removed and there is no any backup of that data left and this is the challenge for service providers to do.

4. *Security Model –*

There is no any fixed and verified security model in the cloud so it is one of the most challenging point in the cloud data security and needs to be corrected.

IV. DATA PROTECTION-AS-A SERVICE

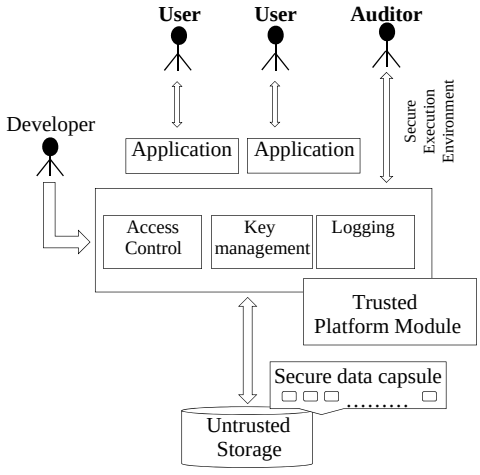

Figure 1.2 Data protection as a service

The above figure 1.2 shows the general model for Data Protection-as-a-Service (DPaaS) [20]. There is a *trusted platform module* (TPM) associated with every server which is responsible for providing security and trust.

This is a high level architecture which combines all possible technologies like access control, encryption, key management, logging, and, checking flow of information which shows DPaaS.

The *secure data capsule* (SDC) is a unit of data which is encrypted and contains various policies related to security. Whatever the data is transmitted by user are confined by this SDC for security.

For avoiding leakage of unauthorized data, *secure execution environment* (SEE) is present which isolates and confines to application execution.

There are various levels in the isolation of inter-SEE and it leads to a potential performance cost because of data marshalling and context switching.

A SEE can be a single virtual machine (VM) or a bunch of virtual machines where a data state is reset every time a data is loaded with a new unit of data.

There are certain requirements in DPaaS approach like-

1. The user which is logged in must be known and who is accessing the services using certain trust policy.
2. The data must be encrypted and then stored into the cloud in order to remove the storage server trust.

A DPaaS model can be best in following four aspects –

1. Online access of data on request of external user when the external user finds the user online.
2. Changes in access controls made by the legitimate users in the cloud.
3. Processing offline various operations like e-mail or the rearrangement of data during changes in schema.
4. Maintenance access like debugging operations by administrator.

V. CLOUD SECURITY STANDARDS

There are three security management standards [21] related to cloud computing – ISO/IEC 27001/02, Information Technology Infrastructure Library (ITIL), and Open Virtualization Format (OVF).

1. *International Organization for Standardization (ISO) 27001/02-*

It describes the essential needs for an Information Security management System (ISMS). It uses ISO 27002 for controlling suitable information security in a ISMS and also is a certification standard.

2. *Information Technology Infrastructure Library (ITIL) –*

It describes the suite of guidelines needed for handling services of information technology. The big advantage of this is that it can be applied to all parts of a system including the operations of cloud. It is a strategic and planned way of handling security issues.

3. *Open Virtualization Format (OVF) –*

It is responsible for an easy software distribution and let the customers chose vendors and services according to their demand. Customers are free to deploy an OVF of any platform.

It is very important for the user to understand the importance of standards setup. A customer before taking a cloud service must go for the above standards vendors and ask them a service level agreements for their service.

A customer also must keep in mind that they rely on OVF to keep tension free themselves and their data also from being threatened. The user must check their service in all respect like portability, efficiency, scalability, integrity, availability etc.

VI. CONCLUSION

Although cloud computing has a potential great features of services like scalability, data on demand and so on, it is still not gaining popularity due to its several vulnerabilities related to the data. Many companies are providing their solutions to protect data but all of them only concentrate on the one stage of data lifecycle while other stages are not covered. Data security is one of the most important point which can help the cloud developers to go ahead. There are various threats to data security which includes both internal as well as external threats. The various types of data like transmission, processing and archived data must be carefully secured. Also, the data security problems like security model, data remanence, data transmission and data migration are some of the inhibiting feature of cloud adoption. DPaaS tries to solve these problems with its features of secure execution environment and secure data capsule. Although DPaaS requires some of pre requisites, it provides a best solution which can erase the data security problems in the cloud and can prove to be the best solution.

REFERENCES

[1] **J. Ari, "PORs: Proofs of Retrievability for Large Files", ACM CCS, pp 584-597, 2007.**
[2] **O. Alina, J. Ari, "Proofs of Retrievability: Theory and Implementation". http://eprint.iacr.org/.**
[3] **O. Alina, J. Ari, "HAIL: A High-Availability and Integrity Layer for Cloud Storage", http://eprint.iacr.org/.**
[4] **W. Brent, S. Hovav, "Compact Proofs of Retrievability", 2008.**
[5] **C. Reza, B. Randal, "Provable Data Possession at Untrusted Stores", CCS 2007.**
[6] **T. Gene, M. Luigi, "Scalable and Efficient Provable Data Possession", 2008.**
[7] **L. Wenjing, W. Cong, "Ensuring Data Storage Security in Cloud Computing", http://www.ece.iit.edu/¬ubisec/IWQoS09.pdf.**
[8] **K. Osama, C. Reza, "MR-PDP: Multiple-Replica Provable Data Possession", 2008.**
[9] **M.Ethan, S. Thomas, "Store, Forget, and Check: Using Algebraic Signatures to Check Remotely Administered Storage", 2006.**
[10] **B. Mike, B. Andrew, "A Cooperative Internet Backup Scheme", 2003.**
[11] **S. Ram, S. Mehul, "Auditing to Keep Online Storage Services Honest", 2007.**
[12] **"Cloud Computing and Security – A Natural Match", April, 2010, http://www.trustedcomputing-group.org/.**
[13] **K. John, "Private Virtual Infrastructure for Cloud Computing".**
[14] **White paper, Oracle, "Information Lifecycle Management for Business Data", June, 2007.**
[15] **Cloud Security Alliance, Security Guidance for Critical Areas of Focus in Cloud Computing V3.0, https://cloudsecurityalliance.org/guidance/csaguide .v3.0.pdf, 2011**
[16] **H. Giles, "Privacy, Security and Indentification in the Cloud", ENISA, 2010.**
[17] **A. Hatem, M. Emam, "Enhanced Data Security Model for Cloud Computing", INFOSO 2012.**
[18] **W. Qioyan, Y. Xiaojun, "A View About Cloud Data Security from Data Life Cycle", IEEE, 2010.**
[19] **W. Xinlei, T. Yubo, "Research of Cloud Computing Data Security Technology", IEEE, 2012.**
[20] **S. Dawn, S. Elaine, "Cloud Data Protection for the Masses", IEEE, 2012.**
[21] **H. Zeljko, P. Kresimir, "Cloud Computing Security Issues and Challenges.", IEEE, 2010.**