International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Security Analysis of Cloud Data Management

***Himanshu Verma*[1], *Mr Rahul Sharma*[2], *Dr Akhil Pandey*[3]**

B-Tech Scholar[1], Professor[2,3]

Department of Computer Science Engineering, Arya College of Engineering & I.T. India, Jaipur

himanshuverma9166@gmail.com [1], *rahulsharma.cs@aryacollege.in*[2], *akhil@aryacollege.in*[3]

ABSTRACT

Cloud computing is a new standard for software and hardware resources according to customers' needs.

This paper includes a brief analysis of data security in a cloud environment across global needs for data protection. The specific research gaps existing in implementation are identified and presented as challenges. The contribution of this review is that it presents an overview of data security analysis, which will be more valuable for data protection in a cloud environment.

Keywords: Cloud computing, Data security

Introduction

Cloud computing is the rescue of computing services over the Internet. Cloud services authorize individuals and businesses to use software and hardware that are managed by third parties at remote locations. Cloud computing realizes computing as a utility. Thus, both users and providers' benefit: providers can reuse their resources, and users acquire and release resources according to their requirements [1]. The cloud provides on-demand self-service in which users can provision resources (network, storage, computing) whenever required without human interaction.

Cloud computing-related technologies include grid computing, utility computing, virtualization, and autonomic computing. Cloud computing is like grid computing in that resources are also coordinated to achieve a common computational objective, but it is one step ahead in that it leverages virtualization technology for better resource utilization and dynamic resource provisioning. Cloud computing acts as a realization of utility computing, which includes on-demand resource provisioning and a utility-based pricing scheme. Cloud computing relates to autonomic computing in a way that it supports autonomic resource provisioning, but its objective is neither to reduce cost nor to reduce system complexity [3].

The cloud environment is up-and-coming and demanding as a new technology, which is an internet-based development. The cloud environment includes public cloud, Private cloud,

Hybrid Cloud, Community Cloud. Cloud services are classified into three types.:

Software as a service (SaaS): Software as a service is software that is accessible via the internet. It contains YouTube, Facebook, and Google apps.

Platform as a Service (PaaS): an operating system, hardware, and network are provided, and the customer installs or develops its own software applications. It includes Amazon DB/S3 [4] and Google App Engine.□ □ Infrastructure as a Service (IaaS): provides just the hardware and network; the customer installs or develops its own operating systems, software, and applications. Examples of IaaS providers include Amazon EC2, □ Geogrid, and Flex Scale [5].

Now the cloud environment revolves around three

□ Functional Units:

Cloud service provider: It is an entity which manages

Cloud Storage Server (CSS) has significant storage space to preserve the clients' data and high computation power.

Client/owner: It is an entity that has large data files to be stored in the cloud and relies on the cloud for data maintenance and computation; it can either be an individual □ consumer or an organization.

User: It is a unit that is registered with the owner and uses the data of the owner stored on the cloud. The user can† be an owner itself as well.

Cloud computing shares characteristics with other computing technologies; hence, it presents unique benefits over other technologies, but at the same time, new security issues arise. According to a 2009 survey from IDCI, 74% of IT managers and CIOs believed that the primary challenge that hindered them from using cloud computing services was cloud computing security. Cloud data storage is built of thousands of cloud storage devices clustered by network, distributed file system, and other storage middleware to provide storage [7]. CDS, however, offers services to assure the integrity of data. But don't provide a solution to the data integrity problem. So, the users themselves must adopt some data security techniques in a cloud environment. Data security in the cloud environment is a major issue that hinders organizations and industries from acquiring cloud services.

Security vulnerabilities always create a fear of data loss and leakage. Data security issues exist in all the service delivery models. Though cloud computing realizes many advantages, including cost-efficiency and flexibility, it requires very little initial setup for starting organizations.

The paper is structured as follows: Section 2 describes data security in a cloud environment; Section 3 discusses research-related studies of previous data security analyses. Key challenges are highlighted in Section 4, and at last we discuss conclusions and future directions.

Data Security in Cloud Environment

Integrity of Data

Data integrity may be easily preserved by employing traditional cryptographic methods such as message authentication codes (MAC). It is a fixed-size block of data based on file F using any secret key. The data owners maintain a small amount of MAC before outsourcing the data, and whenever data is needed, MAC is verified with the previously computed MAC to verify the correctness of the received data from the cloud. In cloud environments, traditional methods are no longer implemented since the data is dynamic and there is huge cloud storage [8]. So, it becomes quite impractical that, to check whether the data is stored securely, we retrieve all the data stored on the server. Integrity threats include data deletion and manipulation. The computation details are not transparent to cloud customers, so the CSP behaves dishonestly and may alter the data.

Data Confidentiality

A user can access the services provided by the SaaS model through a web browser on the internet. So, to protect the data during transmission, HTTPS is implemented. While in IaaS, multiple users' data reside on the same location, in IaaS, confidentiality arises in a way to include isolation over the different user's data.

Data Availability and Management

Availability is affected if the server or service organization is penetrated or spoofed. In the cloud, broad network access (DNS) is one of the main attacks on availability. So, for better service offerings over the internet, users must have reliable DNS.

Authenticity

Cryptographers invent a vast number of primitives for preserving the privacy of users' data. Among these primitives, anonymous password authentication (APA) has been used to ensure the private authentication process. Zero-knowledge authentications in a cloud environment enjoy the benefits of password authentication while offering user privacy preservation.

Data Storage and Maintenance

In a cloud environment, data is dynamically stored over the cloud servers, so the user is unaware of where the data is going to be stored. So, the data locality is an important issue [9]. Another question arises if some investigation occurs: under whose jurisdiction does the investigation occur? The issue can be solved by creating a secure SaaS model that provides reliability for the location of the data. The data in the cloud may become unavailable or effected due to an environmental disaster or server failure. For this, the recovery of important documents must be maintained either by the user on their local disk or by backup services provided by many cloud vendors.

Data Breaches, Leaks and Hacks

Due to the multi-tenancy environment in the cloud, breaching the data will become a potential threat. Data breaches affect two security properties of data: confidentiality, integrity, and authenticity. A data breach may occur internally by some data manager who has direct access to the data or from outside by some malicious hacker. However, confidentiality and integrity issues are addressed by strong cryptographic mechanisms like DES and AES with common PKI infrastructure. Data and key management become an issue for data owners, which can be addressed by combining techniques of attribute-based encryption, proxy re-encryption, and lazy re-encryption.

Data Separation and Filtration

Multi-tenancy is an important characteristic of cloud computing. In multitenancy, multiple users and organizations reside at the same location. So, keeping data separate and maintaining isolation among the users is an important issue [10]. These issues can be solved either by creating a robust virtualization platform or by implementing a Trusted Platform Module embedded on the motherboard.

Literature Review

Virtualization

Jinzhou et al. present a practical architecture to protect data confidentiality for guest virtual machines by constructing a virtualization platform. This is built by hypervisor Xen and trusted platform modules (TPM) integrated on motherboards. Before booting the guest virtual machine, the user prepares an encrypted disk image of the root file system. Then prepare a boot disk image, install Grub on the disk, and put the kernel on it. Send these two disk images to dom0 on the cloud server.

Siefer et al. [12] propose a cyber-physical security framework for data centres that combines security mechanisms in cyber and physical space. The defence framework is based on the time difference between attack detection and the actual attack. Based on physical devices like sensors, it is possible to initiate the defence mechanisms, which include delete, encrypt, and move.

Sijin et al. propose a logging VM for secure logging of auditable file systems. Auditing logs are isolated in another VM on the same host. Through the isolation provided by VMs, the logs are kept safe in another VM even if the working VM gets crashed. So, the privileged user can't access, modify, or delete the data of the unprivileged. Analyse cloud-based virtualization concerns, which include VM security and threats, hypervisor security, data leakage, privacy, data remanence issues in virtualization, and attacks at the virtualization level.

Jasti et al. analyse the security threats that compromise the virtual machine and the hypervisor itself. These threats include VM hoping, VM escape, and mobility. In a multitenancy environment, data from multiple organizations and users resides in the same physical location. In such systems, resources are transparently shared among the VMs of different users. So, a malicious user having control over a VM tries to gain access to another VM or tries to compromise with the data of another VM. For this usage, resources such as C.P.U., memory, and network are analysed on the VM by creating a virtualized environment.

Dimensional Approach

P. Prasad et al. A new way to authenticate using a 3-dimensional approach provides availability of data by overcoming many existing problems like denial of service, data leakage, hacks, and breaches. It also provides more flexibility and capability to meet the new demands of today's complex and diverse network. The idea is that 3D users who want to access the data need to be authenticated to avoid impersonation and data leakage. Now there is a third entity who is either the company's (whose data is stored) employee or customer who wants to access the data; they need to register first, and then before every access to the data, his or her identity is authenticated for authorization.

Information Centric Security

Xiao et al. present a framework to ensure data security in cloud storage systems. SLA is used as the common standard between user and provider, and several technologies are discussed to make data stored in the cloud safe. These technologies include storage protection, transfer protection, and authorization. For secure storage, data is divided into small pieces and saved in different places. If data pieces in one data centre crash, the data can be resumed by the left pieces. Cryptographic protocols like SSL and TLS provide security for communication over networks such as the internet.

Deyan et al. Data security and privacy protection issues in the cloud revolve around the data life cycle, which is from generation to destruction of data. Like in the storage phase, data stored in the cloud needs three security requirements: confidentiality, integrity, and availability. The confidentiality issue is solved by encryption algorithms and key strengths. Similarly, due to the physical characteristics of storage mediums, deleted data still exists and can be retrieved.

Xiaojun et al. describe data security through the data life cycle process. The process includes stages that include creating, storing, using, sharing, archiving, and destructing. Firstly, the client proxy generates data and classifies it by marking. After receiving data, the server checks its integrity. If clients want to use and share data, integrity proof is given by the server.

Zhifeng et al. proposed Auditable MapReduce for the current MapReduce model for making cloud platforms trustworthy. In the map-reduce scenario, all the machines (mappers and reducers) are responsible for performing tasks. Map Function takes key pair input and generates output (intermediate key pair). These intermediate results are sorted by key and taken as input by the reduce function to generate the final output. A-tests are performed and compare output from the worker and the auditor to determine whether the worker is malicious or not. Due to the high overhead in the A-test, it provides P-Accountability since, with a lower number of malicious nodes, there are a smaller number of records to be checked.

A Generic Scheme and Metadata-Based Storage Model

Y. Yang et al. [20] A generic scheme to enable fine-grained data sharing over the cloud that does not require key redistribution or data re-encryption. It is used for attribute-based/predicate encryption and proxy re-encryption.

S. Subashini et al. [21] A metadata-based data segregation and storage methodology and solutions to access this segregated data. In this paper, they investigated the issues of security in data storage in a cloud environment. They proposed

a metadata-based model using which the data residing at the data centre is robbed of its values, and the values are temporarily built-up during runtime and then destroyed once its usage scope is completed. This makes the data invaluable even if an intruder gets access to it. Though this model will take some quantifiable effort to be implemented in real time, it provides a necessary solution for an environment like the cloud, which is showing an adverse potential to become the next-generation enterprise environment. This model, in combination with our multi-tier security model for securing data over transmission, will provide proper crossbars in the wires of malicious users.

Ensure data security Using Cryptographic Protocols

Sood [22] proposes a security model for the whole computing process in the cloud. The characterization and measures are presented for secure storage and efficient retrieval of data from the cloud. Data is classified based on cryptographic parameters like confidentiality, availability, and integrity. Based on the sensitivity rating, data is distributed in public, private, and limited access sections. Data and indexes are encrypted with 128-bit SSL encryption. MAC is generated to check whether the data is not tampered with, and that calculated MAC is matched with the MAC received along with the encrypted file. The model is analysed and tested with the help of the cloud computing simulator Hadoop. The proposed framework is compared with the various security issues that hinder cloud computing from adopting it.

Jian et al. [23] address the issues of privacy problems in cloud computing, like disclosure of sensitive information, which includes personal identification information, unauthorized access to personal data, and protecting the privacy of data when moving outside the organization's boundaries. An anonymity algorithm is proposed in which data is processed by this algorithm before it is sent to the service provider. SP integrates auxiliary information to analyse the data. The anonymity method is different from cryptographic technology; here, the service provider directly uses the data without a key.

Xu et al. [24] analyse the architecture and layers of cloud computing. Analyses the security of the cloud computing platform on which applications are deployed. Design a security framework for the cloud computing platform based on security issues like confidentiality, integrity, availability, nonrepudiation, and reliability. The operation process and module descriptions of the security framework are done.

Wang et al. [25] propose a private matching protocol for matching the personal information to the tuples of the data sets in the data centre. The user checks whether the anonymized data added to the table sets meets K-anonymity. If it meets the K-anonymity data generalization, it is done through a minimal attribute generalization approach to overcome information loss while satisfying the service provider's request and enabling it to send the service to the right client.

Rachal et al. [26] propose an identity management approach (IDM) for data privacy and security that is independent of a trusted third party. The approach makes use of predicates over encrypted data and negotiation when using cloud services. For privacy, an active bundle scheme is implemented, which acts as a middleware agent that includes PII data, privacy policies, and a set of protection mechanisms. It allows the use of IDM applications on untrusted hosts.

CONCLUSION

It is well-known that cloud computing has many potential Lee, "Physical attack protection with human-secure advantages, there are still many actual problems that need to virtualization in data centres," in 42nd International be solved and data are migrating to public or hybrid cloud. Conference on Dependable Systems and Networks According to the analysis for data security, it is expected to Workshop (DSN-W), 2012 IEEE/IFIP, pp. 1-6, have an integrated and comprehensive security solution to [13] Z. Siqi, C. Kang, and Z. Weimin, "Secure Logging Meets the needs of Definitions in depth. for Auditable File System Using Separate Virtual for data security issues, the primary challenges are separation machines," in the IEEE International Symposium on sensitive data and access control. The future objective is to parallel and distribute processing with the design of a set of unified identity management and data security applications, 2009, pp. 153–160