Data Security using Proxy Based Methods in Cloud Computing

Lakshmi K Ramachandrappa [#1], Renukaradhya P C [*2]

[#]*PG Student, CSE Department , Sridevi institute of technology and management, Visveswaraya Technological University*

** Assistant Professor, CSE Department, Sridevi institute of technology and management, Visveswaraya Technological University*

lakshmi9368@gmail.com rpcmtech@gmail.com

Abstract - Data security has consistently been a major issue in information technology. In the cloud computing environment, it becomes particularly serious because the data is located in different places even in all the globe. Data security and privacy protection are the two main factors of user's concerns about the cloud technology. Though many techniques on the topics in cloud computing have been investigated in both academics and industries, data security and privacy protection are becoming more important for the future development of cloud computing technology in government, industry, and business. Data security and privacy protection issues are relevant to both hardware and software in the cloud architecture. This study is to review different security techniques and challenges from both software and hardware aspects for protecting data in the cloud and aims at enhancing the data security and privacy protection for the trustworthy cloud environment. In this paper, we make a comparative research analysis of the existing research work regarding the data security and privacy protection techniques used in the cloud computing.

Keywords — Cloud, Data, Security, Privacy

I. INTRODUCTION

Cloud computing has been envisioned as the next generation paradigm in computation. In the cloud computing environment, both applications and resources are delivered on demand over the Internet as services. Cloud is an environment of the hardware and software resources in the data centers that provide diverse services over the network or the Internet to satisfy user's requirements [1]. The explanation of "cloud computing" from the National Institute of Standards and Technology (NIST) [2] is that cloud computing enables ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. According to the explanation, cloud computing provides a convenient on-demand network access to a shared pool of configurable computing resources. Resources refer to computing applications, network resources, platforms, software services, virtual servers, and computing infrastructure. Cloud computing can be considered as a newcomputing archetype that can provide services on demand at a minimal cost. The three well-known and commonly used service models in the cloud paradigmare software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). In SaaS, software with the related data is deployed by a cloud service provider, and users can use it through the web browsers. In PaaS, a service provider facilitates services to the users with a set of software programs that can solve the specific tasks. In IaaS, the cloud service provider facilitates services to the users with virtual machines and storage to improve their business capabilities. Cloud computing is closely related to but not the same as grid computing [3]. Grid computing integrates diverse resources together and controls the resourceswith the unified operating systems to provide high performance computing services, while cloud computing combines the computing and storage resources controlled by different operating systems to provide services such as large-scaled data storage and high performance computing to users. The overall picture of grid computing has been changed by cloud computing.

II. RELATED WORK

Distribution of data is in a new way of cloud computing comparing with the grid computing. Cloud computing will enable services to be consumed easily on demand. Cloud computing has the characteristics such as on-demand self-service, ubiquitous network access, location independent resource pooling, rapid resource elasticity, usage-based pricing, and transference of risk. These merits of cloud computing have attracted substantial interests from both the

industrial world and the academic research world.Cloud computing technology is currently changing the way to do business in theworld. Cloud computing is very promising for the IT applications; however, there are still some problems to be solved for personal users and enterprises to store data and deploy applications in the cloud computing environment.One of the most significant barriers to adoption is data security, which is accompanied by issues including compliance, privacy, trust, and legal matters [4, 5]. The role of institutions and institutional evolution is close to privacy and security in cloud computing [6]. Data security has consistently been a major issue in IT. Data security becomes particularly serious in the cloud computing environment, because data are scattered in different machines and storage devices including servers, PCs, and various mobile devices such as wireless sensor networks and smart phones. Data security in the cloud computing is more complicated than data security in the traditional information systems. To make the cloud computing be adopted by users and enterprise, the security concerns of users should be rectified first tomake cloud environment trustworthy.The trustworthy environment is the basic prerequisite to win confidence of users to adopt such a technology. Latif et al. discussed the assessment of cloud computing risks [7]. Before thedata security issues arediscussed, the functions of cloud computing are analyzed first. Cloud computing is also known as on-demand service. In the cloud computing environment, there is a cloud service provider that facilitates services and manages the services. The cloud provider facilitates all the services over the Internet, while end users use services for satisfying their business needs and then pay the service provider accordingly. Cloud computing environment provides two basic types of functions: *computing* and *data storage.* In the cloud computing environment, consumers of cloud services do not need anything and they can get access to their data and finish their computing tasks just through the Internet connectivity. During the access to the data and computing, the clients do not even know where the data are stored and which machines execute the computing tasks. Coming to data storage, data protection and security are the primary factors for gaining user's trust and making the cloud technology successfully used. A number of data protections and data security techniques have been proposed in the research field of cloud computing. However, data protection related techniques need to be further enhanced. Services of cloud computing are provided across the entire computing spectrum. Nowadays, organizations and companies are moving and extending their business by adopting the cloud computing to lower their cost. This can contribute to free more man-powers to focus on creating strategic differentiation and business division of labor is clearer. The cloud is growing continuously because it could provide high performance computational services at cheaper rates. Famous IT companies such as Microsoft (http://azure.microsoft.com/), Amazon (http://aws.amazon .com/), Google (https://cloud.google .com/), and Rakespace (http://www.rackspace.com/) have provided cloud service on the Internet. The concept of cloud has a number of implementations based on the services from service providers. For example, Google Apps Engine, Microsoft Azure, and Amazon Stack are popular implementations of cloud computing provided by cloud service providers, that is, Google, Microsoft, and Amazon companies. Besides, the ACME enterprise implemented VMware based v-Cloud for permitting multiple organizations to share computing resources. According to the difference of access scope, cloud can be divided into three types: *public cloud*, *private cloud*, and *hybrid cloud.* Public cloud is as the property of service provider and can be used in public, private cloud refers to being the property of a company, and hybrid cloud is the blends of public and private cloud.Most of the existing cloud services are provided by large cloud service companies such as Google, Amazon, and IBM. A private cloud is a cloud in which only the authorized users can access the services from the provider. In the pubic cloud anybody can use the cloud services whereas the hybrid cloud contains the concept of both public and private clouds. Cloud computing can save an organization's time and money, but trusting the system is more important because the real asset of any organization is the data which they share in the cloud to use the needed services by putting it either directly in the relational database or eventually in a relational database through an application.

III. DATA INTEGRITY

Data integrity is one of the most critical elements in any information system. Generally, data integrity means protecting data from unauthorized deletion, modification, or fabrication. Managing entity's admittance and rights to specific enterprise resources ensures that valuable data and services are not abused, misappropriated, or stolen. Data integrity is easily achieved in a standalone system with a single database. Data integrity in the standalone system is maintained via database constraints and transactions, which is usually finished by a database management system (DBMS). Transactions should follow ACID (atomicity, consistency, isolation, and durability) properties to ensure data integrity.Most databases support ACID transactions and can preserve data integrity. Authorization is used to control the access of data. It is the mechanism by which a system determines what level of access a particular authenticated

user should have to secure resources controlled by the system. Data integrity in the cloud system means preserving information integrity.The data should not be lost or modified by unauthorized users. Data integrity is the basis to provide cloud computing service such as SaaS, PaaS, and IaaS. Besides data storage of large-scaled data, cloud computing environment usually provides data processing service. Data integrity can be obtained by techniques such as RAID-like strategies and digital signature. Owing to the large quantity of entities and access points in a cloud environment, authorization is crucial in assuring that only authorized entities can interact with data. By avoiding the unauthorized access, organizations can achieve greater confidence in data integrity. The monitoring mechanisms offer the greater visibility into determining who or what may have altered data or system information, potentially affecting their integrity. Cloud computing providers are trusted to maintain data integrity and accuracy.However, it is necessary to build the third party supervisionmechanism besides users and cloud service providers. Verifying the integrity of data in the cloud remotely is the perquisite to deploy applications. Bowers et al. proposed a theoretical framework "Proofs of Retrievability" to realize the remote data integrity checking by combining error correction code and spot-checking [10]. The HAIL system uses POR mechanism to check the storage of data in different clouds, and it can ensure the redundancy of different copies and realize the availability and integrity checking [11]. Schiffman et al. proposed trusted platform module (TPM) remote checking to check the data integrity remotely [12].

IV. DATA

Data availability means the following: when accidents such as hard disk damage, IDC fire, and network failures occur, the extent that user's data can be used or recovered and how the users verify their data by techniques rather than depending on the credit guarantee by the cloud service provider alone. The issue of storing data over the transboarder servers is a serious concern of clients because the cloud vendors are governed by the local laws and, therefore, the cloud clients should be cognizant of those laws.Moreover, the cloud service provider should ensure the data security, particularly data confidentiality and integrity. The cloud provider should share all such concerns with the client and build trust relationship in this connection. The cloud vendor should provide guarantees of data safety and explain jurisdiction of local laws to the clients. The main focus of the paper is on those data issues and challenges which are associated with data storage location and its relocation, cost, availability, and security. Locating data can help users to increase their trust on the cloud. Cloud storage provides the transparent storage service for users, which can decrease the complexity of cloud, but it also decreases the control ability on data storage of users. Benson et al. studied the proofs of geographic replication and succeeded in locating the data stored in Amazon cloud [14].

4.1. Reliable Storage Agreement. Themost common abnormal behavior of untrusted storage is that the cloud service providers may discard part of the user's update data, which is hard to be checked by only depending on the simple data encryption. Additionally, a good storage agreement needs to support concurrent modification by multiple users. Mahajan et al. proposed Depot which can guarantee Fork-Join-Causal-Consistency and eventual consistency. It can effectively resist attacks such as discarding and it can support the implementation of other safety protections in the trusted cloud storage environment (such as Amazon S3). Feldman et al. proposed SPORC [8], which can implement the safe and reliable real-time interaction and collaboration for multiple users with the help of the trusted cloud environment, and untrusted cloud servers can only access the encrypted data. However, operation types supported by reliable storage protocol support are limited, andmost of the calculations can only occur in the client.

4.2. *Reliability of Hard-Drive.* Hard-drive is currently the main storage media in the cloud environment. Reliability of hard disks formulates the foundation of cloud storage. Pinheiro et al. studied the error rate of hard-drives based on the historical data of hard-drive [9]. They found that the error rate of hard-drives is not closely relevant to the temperature and the frequency to be used, while the error rate of hard-drives has the strong clustering characteristics. Current SMARTmechanism could not predict the error rate of hard disks. Tsai et al. studied the correlation between the soft error and hard error of hard disks, and they also found that the soft error could not predict the hard errors of harddrives precisely [13], only about 1/3 probability that hard errors follow the soft errors.

V. DATA PRIVACY

Privacy is the ability of an individual or group to seclude themselves or information about themselves and thereby reveal them selectively [12]. Privacy has the following elements. (i) When: a subject may be more concerned about the current or future information being revealed than

information from the past. (ii) How: a user may be comfortable if his/her friends can manually request his/her information, but the user may not like alerts to be sent automatically and frequently.
(iii) Extent: a user may rather have his/her information reported as an ambiguous region rather than a precise point.
In commerce, consumer's context and privacy need to be protected and used appropriately. In organizations, privacy entails the application of laws, mechanisms, standards, and processes by which personally identifiable information is managed [6]. In the cloud, the privacy means when users visit the sensitive data, the cloud services can prevent potential adversary from inferring the user's behavior by the user's visit model (not direct data leakage). Researchers have focused on Oblivious RAM (ORAM) technology. ORAM technology visits several copies of data to hide the real visiting aims of users. ORAM has been widely used in software protection and has been used in protecting the privacy in the cloud as a promising technology. Stefanov et al. proposed that a path ORAM algorithm is state-of-the-art implementation [7]. The privacy issues differ according to different cloud scenarios and can be divided into four subcategories as follows:
(i) how to enable users to have control over their data when the data are stored and processed in cloud and avoid theft, nefarious use, and unauthorized resale,
(ii) how to guarantee data replications in a jurisdiction and consistent state, where replicating user data to multiple suitable locations is an usual choice, and avoid data loss, leakage, and unauthorized modification or fabrication,
(iii) which party is responsible for ensuring legal requirements for personal information,
(iv) to what extent cloud subcontractors are involved in processing which can be properly identified, checked, and ascertained.

5.1. Service Abuse. Service abuse means that attackers can abuse the cloud service and acquire extra data or destroy the interests of other users. User data may be abused by other users. Deduplication technology has been widely used in the cloud storage, which means that the same data often were stored once but shared by multiple different users. Thiswill reduce the storage space and cut down the cost of cloud service providers, but attackers can access the data by knowing the hash code of the stored files. Then, it is possible to leak the sensitive data in the cloud. So proof of ownership approach has been proposed to check the authentication of cloud users [48]. Attackers may lead to the cost increase of cloud service. Fraudulent resource consumption is a kind of attack on the payment for cloud service. Attackers canconsume the specific data to increase the cost for cloud service payment. Idziorek et al. proposed this question and researched on the detection and identification of fraud resource consumption.

5.2. Averting Attacks. The cloud computing facilitates huge amount of shared resources on the Internet. Cloud systems should be capable of avertingDenial of Service (DoS) attacks. Shen et al. analyzed requirement of security services in cloud computing. The authors suggest integrating cloud services for trusted computing platform (TCP) and trusted platform support services (TSS). The trusted model should bear characteristics of confidentiality, dynamically building trust domains and dynamic of the services. Cloud infrastructures require that user transfers their data into cloudmerely based on trust. Neisse et al. analyzed indifferent attacks scenarios on Xen cloud platform to evaluate cloud services based on trust. Security of data and trust in cloud computing is the key point for its broader adoption. Yeluri et al. focused on the cloud services from security point of view and explored security challenges in cloud when deploying the services. Identity management, data recovery and management, security in cloud confidentiality, trust, visibility, and application architecture are the key points for ensuring security in cloud computing.

CONCLUSION

Cloud computing is a promising and emerging technology for the next generation of IT applications. The barrier and hurdles toward the rapid growth of cloud computing are data security and privacy issues. Reducing data storage and processing cost is a mandatory requirement of any organization, while analysis of data and information is always the most important tasks in all the organizations for decision making. So no organizations will transfer their data or information to the cloud until the trust is built between the cloud service providers and consumers. A number of techniques have been proposed by researchers for data protection and to attain highest level of data security in the cloud. However, there are still many gaps to be filled by making these techniques more effective. More work is required in the area of cloud computing tomake it acceptable by the cloud service consumers. This paper surveyed different techniques about data security and privacy, focusing on the data storage and use in the cloud, for data protection in the cloud computing environments to build trust between cloud service providers and consumers.

REFERENCES

[1] Neal Leavitt, " Is cloud computing really ready for prime time?," in IEEE Computer Socciety, 2009.

[2] W. A. Jansen, "Cloud Hooks: Security and Privacy Issues in Cloud Computing," in 44th Hawaii International Conference on System Sciences, pp. 1–10, Koloa, Hawaii, January 2011.

[3] W. Itani, A Kayssi, and A Chehab, "Privacy as a service: Privacy-aware data storage and processing in cloud computing architectures, " in IEEE conference on Dependable, Autonomic and Secure Computing, DASC '09, pp 711–716, December 2009.

[4] M. Jensen, J. Schwenk, N. Gruschka, and L.L. Iacono, "On technical security issues in cloud computing," in IEEE International Conference on Cloud Computing, CLOUD '09, pp. 109–116, September 2009.

[5] Meiko Jensen, Nils Gruschka, and Ralph HerkenhÃner. A, " A survey of attacks on web services," in Journal of Computer Science - Research and Development, pp. 185–197, 2009.

[6] Ronald L. Krutz and Russell Dean Vines, "Cloud Security: A Comprehensive Guide to Secure Cloud Computing," Wiley Publishing, 2010.

[7] Rongxing Lu, Xiaodong Lin, Xiaohui Liang, and Xuemin Shen, "Secure provenance: The essential of bread and butter of data forensics in cloud computing, " in Proceedings of 5th ACM Symposium on Information, Computer and Communications Security, ASIACCS '10, pp. 282–292, 2010 Peter Mell and Timothy Grance. The nist definition of cloud computing. Technical Report 800-145, National Institute of Standards and Technology (NIST), Gaithersburg, MD, September 2011.

[8] S. Subashini and V. Kavitha, " Review: A survey on security issues in service delivery models of cloud computing, " in Journal of Network and Computer Applications,pp. 1–11, January 2011.

[9] Cong Wang, Qian Wang, Kui Ren, and Wenjing Lou, " Ensuring data storage security in cloud computing," in 17th Internation Workshop on Quality of Service, IWQoS, pp 1–9, July 2009.

[10] Cong Wang, Qian Wang, Kui Ren, and Wenjing Lou, "Privacypreserving public auditing for data storage security in cloud computing," in Proceedings IEEE INFOCOM, pp 1–9, March 2010.

[11] Ardagna, Danilo, Giuliano Casale, Michele Ciavotta, Juan F Perez and Weikun Wang, Quality-of-service in cloud computing:modelling techniquesand their applications," in Journal of Internet Services and Applications, 2014.

[12] Alan T Litchfield and Jacqui Althouse, "A systematic review of cloud computing, big data and databases on the cloud," in Proceedings of the Americas Conference on Information Systems, pp 1–19, 2014.

[13] Monjur Ahmed, Alan T Litchfield and Chandan Sharma, "A distributed security model for cloud computing," in Proceedings of the Americas Conference on Information Systems, 2016.

[14] Z. Masetic, K. Hajdarevic, N. Dogru. Cloud Computing Threats Classification Model Based on the Detection Feasibility of Machine Learning Algorithms," in 40th International Conference on Information and Technology, Electronics and Microelectronics, 2017.