A Study On Secure Data Storage In Public Clouds

T. Dhanur Bavidhira
PG Scholar: dept. of Computer Science and Engineering
Velalar College of Engineering and Technology
Erode, India
e-mail: dhanur14@gmail.com

C. Selvi
Asst. Professor: dept. of Computer Science and Engineering
Velalar College of Engineering and Technology
Erode, India
e-mail: selviilango.cse@gmail.com

***Abstract*—** This paper focuses on the study of various existing cloud storage mechanisms with their related security frameworks for realizing the efficient cloud storage in a secured environment. The key feature for the growing popularity of cloud computing relies on the efficient management of stored data in a highly secure way for remote accessing. Ensuring the integrity and availability of user's data stored in the cloud has always been an important aspect for its quality of service. While storing data in cloud, lots of issues with respect to security are being cropping out as clients have no direct physical control over their outsourced data. In addition, their vulnerabilities to external threats are increasing as cloud provides storage and accessing services via world-wide domain networking. This study will help in identifying different performance measures for secure available of data in cloud storage mechanisms.

Keywords- *Efficient storage, security, data reliability, availability, storage techniques, encryption techniques.*

*****

I. INTRODUCTION

Over the past few decades, the dependency on the use of computer has been increased tremendously for remote data storage with security. Therefore, cloud computing has paved the way for realizing a secured data storage in a remote online location. Cloud computing is a computing technology, that shares both hardware and software resources via internet to users. This becomes the most alluring technology due to its economical benefit with high degree of flexibility. According to NIST, "Cloud computing is a model for enabling 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" [15]. The concept of virtualization is used in the cloud computing to run a variety of applications concurrently. Cloud saves the user's data in an off-site storage system instead of storing it on user's local storage devices that enables the user to access their data from remote location through internet. There is an extensive scope of cloud computing as many of the organizations have adopted it. Use of cloud computing in organizations can increase the capacity and capabilities of the software by many folds. Like organizations many people are moving towards cloud computing because it allows user to access data and resources from any geographical location at any time. It also reduces infrastructure costs, scalability and maintenance cost. Here the resources are provided based on the demand and billed based on pay per use. Cloud offers three main types of services they are Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as Service (IaaS) [15]. The most common type of service offered by cloud is Storage as a Service (e.g. Amazon S3, Google Drive, Dropbox and OneDrive) where data are stored in data centers and can be accessed when it is required. Security and cost are the major threat factors that are being associated in storage. Security is one of the primary concerns in adopting cloud. Day by day, the data in cloud is getting increased and some methods are needed to guarantee security for the stored data. During data communication over cloud unauthorized users may hack the data this alarms users data in threat. Users from various enterprises with various privilege levels often interact with computing resources. This may lead to accidental misuse of data and have to be monitored for its abuse. However, several recent surveys [1], [2] show that 88% of cloud consumers are worried about their data security. In this survey, a brief study on current secure data storage mechanisms in cloud has been presented.

II. LITERATURE REVIEW

Huaqun Wang, et al.,[11] has paper proposed a new security system that uses the idea of identity-based proxy-oriented data uploading and remote data integrity checking in public cloud (ID-PUIC). This system is mainly built with an objective of solving new security problems for efficient processing of clients data in public cloud when there is an increase in number of clients to PCS (public cloud servers). It employs ID-PUIC protocol using bilinear pairings technique. The ID-PUIC protocol is provably secure and efficient than the computational Diffie-Hellman problem. Based on the original client's authorization the proposed system can perform private remote data integrity checking, delegated remote data integrity checking and public remote data integrity checking.

Xuefeng Liu and Yuqing Zhang [3] have proposed a system for dynamic groups in the cloud called Mona, a secure multi owner data sharing scheme. Any user in the cloud can share their data anonymously with others by leveraging group signature and dynamic broadcast encryption techniques. The computation cost for encryption and the storage overhead are independent with the number of revoked users.

Garima and Naveen [10], "Triple Security of Data in Cloud Computing", proposed a system for enhancing security in cloud by applying three algorithms namely: DSA (Digital Signature Algorithm), AES (Advanced Encryption Standard) and Steganography. For data encryption, DSA is done for authentication followed by AES for encryption and finally applied Steganography for concealing data within audio file to have maximum security. By applying the algorithms in the reverse order the receiver can decrypt the data but the issues found here is high complexity in time since the algorithms are applied one after the other.

In Cong Wang & Kui Ren [4] Privacy-Preserving Public Auditing for Secure Cloud Storage has mainly focuses on third-party auditor (TPA) to perform audits for numerous users concurrently and efficiently. Hence to achieve maximum security the third-party auditor should not bring any additional vulnerability to the system or extra burden for users. The performance analysis of the proposed schemes proved that they are highly secure and efficient.

The sharing of cloud computing resources like storage, services and applications with other tenants is very risky as they accidentally get other tenants information. Multi-tenancies are considered as an important feature in cloud computing resource utilization. Hence Mohamed Al Morsy et al., [5] has delivered a system that secures multi-tenancy by the segregation of cloud users.

Sanjoli and Jasmeet [8], "Cloud data security using authentication and encryption technique", It proposes the combination of two different algorithms Extensible Authentication Protocol (EAP) - Challenge-Handshake Authentication Protocol (CHAP) and Rijndael Encryption Algorithm for enabling high security. The former algorithm is applied for authentication purpose and the later for data encryption. Rijndael Encryption Algorithm makes the system more secure. In this paper Client side security has been concentrated for attaining high data security.

According to Younis et al., [6] has discussed about giving a secure Multi-tenancy in the cloud computing that requires isolation among customer's data. In cloud computing data are stored in different countries that face various regulations and legal systems. Securing of customers data is the critical factor in cloud computing because the sharing of cloud computing resources may give arise to unauthorized access.

Shirole and Sanjay [9], "Data Confidentiality in Cloud Computing with Blowfish Algorithm", proposes a system with OTP (One-Time Password) for authentication and Blowfish algorithm for encryption. This provides reliability and easiness in storing secure data. Here ciphered data is uploaded in the cloud by applying encryption on the plain text and whenever the data is needed, it is obtained from the cloud and is stored on the system in the plain format. Hence this protects the data internally.

According to Keiko David, Eduardo and Eduardo [7], "An analysis of security issues for cloud computing", this paper discusses the various threat areas that require high level of security. External data storage, using public internet for communication, multi-tenancy, data integration are some of the major risk areas that has been dealt in detail.

III. Secure Storage Techniques in Cloud Computing

The different kinds of storage techniques for secure cloud storage are discussed below.

A. Authentication based on Identity

In cloud computing resources and services are between various users and there arises critical issues in terms of security. Hence authentication has become an important factor for both users and the service providers. Once Secure Socket Layer Authentication Protocol (SAP) is applied in cloud computing it becomes very complex in terms of communication and computation overhead which results in poor efficiency for both clients and Cloud services. To overcome this identity based authentication protocol is used. On comparing, identity based authentication protocol is proved to be more efficient and weightless. In this paper [16], identity-based hierarchical model for cloud computing and its corresponding encryption and signature schemes are used to have security in cloud communication.

B. Implicit Storage Security to Online Data

The primary focus on cloud computing is securing the data in the distributed servers. Securing data storage in cloud computing can be done in two ways either explicitly or implicitly. Explicit method of securing data is the conventional approach to store data. Here data is stored and data backup resides on single server that permits clients to access data based on the use of passwords. But this architecture proves to be inefficient for securing data in many applications. An alternate way to this conventional approach is implicitly securing data. [17] It uses the idea of data partitioning where the partitions are stored on the servers aimlessly and does not need to be encrypted. Reconstruction of data needs to access each and every server to recreate the client's original data. It is important

to have the information where the partitions are stored on server and must back it up on a single server. Hence data partition must be done in such a way that the knowledge of each partitioned piece should be indebted to recreate the data. It is helpful to employ implicit security for online data storage in a cloud computing environment.

C. Public Auditing for Dynamic Data Storage Security

The main concern in cloud storage is authentication for data integrity in the cloud servers. There is always a need for dynamic data operation and public auditability to ensure the remote data integrity. [18] To attain data dynamics, improvised Merkle Hash Tree (MHT) model is used for block tag authentication. For managing the several auditing task efficiently the technique of bilinear aggregate signature is used. Thus TPA will be capable of performing multiple auditing tasks at a time in multiuser setting. This scheme is extremely proficient and provably secure.

D. Efficient Third Party Auditing

In Cloud computing users can store their data in a remote location and can be accessed from anywhere for various services. The major issue faced by the user is he unsure about the data in cloud possesses integrity or not. Third party auditing can employ symmetric key cryptography technique to carry out the auditing without having any local copies of data. This greatly helps in the reduction of computation and transmission loads. [19]In order to accomplish data storage security data blocks have to be signed in prior to data outsourcing to cloud. For this BLS (Boneh–Lynn–Shacham) algorithm is used due to its efficiency. TPA can perform multiple auditing tasks for various simultaneously.

E. Effective and Secure Storage Protocol

Nowadays more and more users outsource their data are to cloud providers who offer sufficient storage with minimum cost in order to relieve from the local storage burden. However the users face an important risk towards integrity and confidentiality of data security. Hence a new security protocol has designed to overcome this. [20] A protocol is built in combination with Elliptic curve cryptography and Sobol Sequence. It is exercised to verify the data integrity arbitrarily without getting original data. Communication and I/O costs can be brought down considerably creating probabilistic proofs for a random set of blocks from the cloud server.

F. Dynamic Data Storage in Cloud

It is not an easy job to maintain all data in the cloud securely. In cloud data storage is not fully trustable as clients do not have any local copy of data. To avoid this issue a new system is designed that employs data reading algorithm for verifying integrity of the data [21]. In future the data can be obtained by using automatic data reading algorithm. Integrity of data can be checked before and after insertion of data to the file.

G. Data Storage Security

Cloud Computing gives an opportunity to share their services and resources to an open environment through network. This leads to many troubles in security because transmission of data on network is susceptible to attack. To avoid this data have to be encrypted in cloud environment. A method [12] has been constructed that has three backup sites to recover the data after disaster. These backup sites are situated at remote location from the main server. This method employs three algorithms namely SHA algorithm for encryption, GZIP algorithm for compression and SFSPL algorithm for splitting files. It gives a cross platform model for secure communication.

H. Secure and Dependable Storage

Storage in cloud computing allows users store their data remotely without worrying about the local data management. However there are some issues concerning to the correctness of data as the user doesn't possess any local copy of data. This situation leads to some serious issues in determining the reliability of user's data because the users do not have a copy of their record. A new system [22] is introduced that helps the client to examine data storage in cloud. The system relies on Homomorphic token with Reed-Solomon erasure correcting code technique for identifying of server misbehaving and ensures the correctness insurance. The users can delegate the task to third party auditor when they don't have enough time and resources. This technique also permits users to perform dynamic operations like insertion, deletion and update on stored data in a remote location.

I. Optimal Cloud Storage

Optimal cloud storage is an efficient feature for the resource providers to store the data in the cloud. A taxonomic approach is adopted to attain optimal cloud storage service. A new prototype [23] called NubiSave is proposed to have optimal storage in cloud. This prototype has an architecture that serves as blueprint for optimal storage controller. NubiSave is freely available and has to be integrated with front ends for future work. The proposed architecture consists of three components namely data processor, data verifier and token generator. Data processor helps in data processing before it is being sent to the cloud. Data verifier validates the data whether it has been tampered or not in the cloud. Token generator is responsible for generating the tokens and assists the storage providers for retrieving the parts of client's data.

J. Accessing and Storing Small Files

Hadoop Distributed File System (HDFS) is built to hold up the internet services comprehensively. [24] This system is capable of accessing the applications having large data sets in high-throughput but it fails to do so in small files. The main

issues that are addressed in the small files are burden on the NameNode of HDFS experiences huge load made by large numbers of small files, correspondence between small files are not taken into account for data placement and no optimization mechanism is present. A novel approach is designed to enhance the storage of small file in HDFS. It uses cut off points to analyze the size of the file and based on the correspondence level between the file, the files are divided into three type namely structurally-related files, logically-related files and independent files. Finally optimization technique is employed base on the type of file.

K. File Storage and Security Fortification

A security structure for file storage has been introduced that uses idea of master and slave servers. This eradicates the direct communication between the client and the cloud server. In this proposed model [23] master server is in charge of processing the clients request and the slave server is in charge of chunking the files. Chunking operation is done to store duplicates of records and backup for recovery. It uses token generation and merging algorithms to attain storage integrity and data availability. It allows users to do powerful data dynamic operations.

L. Secure Public Auditing for Privacy Preserving

In data outsourcing, cloud servers have plenty of control on the outsourced data than the data owners. To check the integrity of the client's data in the cloud server Third Party Auditor (TPA) is introduced [14]. The users delegate the control to Third Party Auditor to verify the data integrity periodically to reduce their overhead. In Third Party Auditor security is not compromised since user's data are not given straightforwardly. It uses Elliptic curve Cryptography (ECC) to give quick processing and computation for cloud storage servers. It is also capable of doing multiple auditing tasks.

IV. DATA ENCRYPTION TECHNIQUES

TABLE I. COMPARATIVE ANALYSIS ON VARIOUS EXISTING DATA ENCRYPTION TECHNIQUES

Encryption Techniques	**Key Length**	**Computation time (in ms)**	**Merits**	**Demerits**
AES	2048	2390	Protection from tampering and brutal force attack.	Less efficient for privacy protection
Caesar cipher and Vigenere Cipher	Identical to size given alphabet	1060	Provides triple encryption	Less efficient.
Craig	poly(d)	O(n^2)	Good in	• Fully homomorphic • High computation
Gentry construct homomorphism encryption scheme			security aspect	
Hierarchical Identity Based Encryption (HIBE)	O(t)	Computation time grows linearly in the depth `l ` of the hierarchy	• Highly efficient • Low computing overhead	• Collusion • Resistance • User accountability
Blowfish	1024	500	• High computation speed • Good Key strength.	• Requires more memory • Relatively long key setup time
RC4	1024	169	Simple and fast	Not efficient for block ciphering.
RSA-1,2	1024,2048	5199	Public Key encryption.	Lower Speed

V. CONCLUSION

A study on various security challenges in different cloud computing platforms has been discussed and highlighted. It also provides a comparative model on different existing data encryption techniques used in the cloud based on their merits and demerits. In real-time, some security features are difficult to realize and hence further enhancements have to take place to improve the efficiency of algorithms to attain the high threshold in secure data storage in cloud computing.

REFERENCES

[1] Lan Zhou, Vijay Varadharajan and Michael Hitchens, "Achieving Secure Role-Based Access Control on Encrypted Data in Cloud Storage", IEEE Transactions on Information Forensics and Security, vol. 8, no. 12, 2013.

[2] Manachai Toahchoodee, Ramadan Abdunabi, Indrakshi Ray and Indrajit Ray, "A Trust-Based Access Control Model for Pervasive Computing Applications", IFIP Annual Conference on Data and Applications Security and Privacy, 2009.

[3] Xuefeng Liu, Yuqing Zhang, Boyang Wang and Jingbo yan, "Mona: Secure Multi-Owner Data Sharing for Dynamic Groups in the Cloud", IEEE Transactions on Parallel and Distributed System, vol. 24, no. 6, 2013.

[4] Cong Wang, Sherman S.M. Chow, Qian Wang, Kui Ren and Wenjing Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage", IEEE transactions on computers, vol. 62, no. 2, 2013.

[5] Younis A.Younis, Madjid Merabti and Kashif Kifayat, "Secure Cloud Computing for Critical Infrastructure: A Survey", International Journal of Computer Applications, 2013.
[6] Thapliyal, Meenakshi, Hardwari Lal Mandoria, and Neha Garg, "Data Security Analysis in Cloud Environment: A Review", International Journal of Innovations & Advancement in Computer Science, vol. 2, no. 1, 2014.
[7] Keiko Hashizume, David G Rosado, Eduardo Fernández-Medina and Eduardo B Fernandez, "An Analysis of Security Issues for Cloud Computing", Journal of Internet Services and Applications, 2013.
[8] Singla and Jasmeet Singh, "Cloud Data Security using Authentication and Encryption Technique", Global Journal of Computer Science and Technology, 2013.
[9] Subhash and Shirole Bajirao, "Data Confidentiality in Cloud Computing with Blowfish Algorithm", International Journal of Emerging Trends in Science and Technology, 2014.
[10] Saini, Garima and Naveen Sharma, "Triple Security of Data in Cloud Computing", International Journal of Computer Science & Information Technologies, 2014.
[11] Huaqun Wang, Debiao He and Shaohua Tang, "Identity-Based Proxy-Oriented Data Uploading and Remote Data Integrity Checking in Public Cloud", IEEE Transactions on Information Forensics and Security, vol. 11, no. 6, 2016.
[12] S.Sajithabanu and Dr.E.George Prakash Raj, "Data Storage Security in Cloud", International Journal of Computer Science and Technology, vol. 2, no. 4, 2011.
[13] http://crypto.stanford.edu/~dabo/papers/shibe.pdf.
[14] Cong Wang, Sherman S.M.Chow, Qian Wang, Kui Ren and Wenjing Lou, "Privacy Preserving Public Auditing for Secure Cloud Storage", IEEE Transactions on Computers, vol. 62, no.2, 2013.
[15] https://www.nist.gov/itl/cloud-computing.
[16] Hongwei Li, Yuanshun Dai, Ling Tian and Haomiao Yang, "Identity-Based Authentication for Cloud Computing", First International Conference on Cloud Computing, no.2, pp. 157–166, 2009.
[17] Abhishek Parakh and Subhash Kak, "Online Data Storage using Implicit Security", Information Sciences, vol. 179, 2009.
[18] Wang Q, Wang C, Kui Ren and Wenjing Lou, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing", IEEE Transactions on Parallel and Distributed Systems, vol.22, no.5, 2011.
[19] Balakrishnan.S, Saranya.G, Shobana.S and Karthikeyan.S, "Introducing Effective Third Party Auditing (TPA) for Data Storage Security in Cloud", International Journal of Computer Science and Technology, vol 2, no.2, 2011.
[20] Syam Kumar P and Subramanian R, "An Efficient and Secure Protocol for Ensuring Data Storage security in Cloud Computing", International Journal of Computer Science, vol.8, no.1, 2011.
[21] Kumar S P, Subramanian R, "An efficient and secure protocol for ensuring data storage security in Cloud Computing", International Journal of Computer Science Issues, vol. 8, no. 1, 2011.
[22] Wang C, Wang Q, Kui Ren, Ning Cao and Wenjing Lou, "Toward Secure and Dependable Storage Services in Cloud Computing", IEEE Transactions on Services Computing, vol.5, no.2, 2012.
[23] Punyada M. Deshmukh, Achyut S. Gughane, Priyanka L. Hasija and Supriya P. Katpale "Maintaining File Storage Security in Cloud Computing", International Journal of Emerging Technology and Advanced Engineering, vol. 2, no.10, 2012.
[24] Bo Dong, Qinghua Zheng, Feng Tian, Kuo-Ming Chao, Rui Ma and Rachid Anane, "An Optimized Approach for Storing and Accessing Small Files on Cloud Storage", Journal of Network and Computer Applications, 2012.