To Improve Data Storage Security Levels in the Cloud

Kolluru Venkata Nagendra
Assistant Professor,
Department of CSE
Geethanjali Institute of Science and Technology.
Gangavaram,Nellore, Andhra Pradesh, India
kvnscholar@gmail.com

N. Haritha,
Assistant Professor,
Department of Computer Science,
Sriharsha Institute of PG Studies,
Nellore, Andhra Pradesh, India
nharithamca@gmail.com

Abstract: Now-a-Days Cloud Computing is an emerging technology, that works on the principle of pay-per-use. It offers services like Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), Infrastructure-as-a-Service (IaaS), Storage as a Service and many more. Cloud computing used for database and software applications to centralize the data. All the cloud users openly store their data on the cloud service provider's service centers. Here, the management of data and services are not fully confidential. So that the security of the cloud stored data becomes an open challenging task in the field of Cloud Computing. To increase the levels of data security in the cloud, this paper deals with some techniques like public auditability, Homomorphism Linear Authenticator (HLA).

Keywords: *Cloud Computing, Data Security, public auditability, Homomorphism Linear Authenticator (HLA)*

*****

I. INTRODUCTION

Cloud computing is a recent trending in IT that where computing and data storage is done in data centers rather than personal portable PC's. It refers to applications delivered as services over the internet as well as to the cloud infrastructure – namely the hardware and system software in data centers that provide this service. The sharing of resources reduces the cost to individuals. The best definition for Cloud is defined in as large pool of easily accessible and virtualized resources which can be dynamically reconfigured to adjust a variable load, allowing also for optimum scale utilization. The key driving forces behind cloud computing is the ubiquity of broadband and wireless networking, falling storage costs, and progressive improvements in Internet computing software. The main technical supporting of cloud computing infrastructures and services include virtualization, service-oriented software, grid computing technologies, management of large facilities, and power efficiency. The key features of the cloud are agility, cost, device and location independence, multi tenancy, reliability, scalability, maintenance etc.

The cloud can be deployed in three models. They are described in different ways. In generalized it is described as below:

A. Public Cloud:

Public cloud describes cloud computing in the traditional mainstream sense, whereby resources are dynamically provisioned on a fine-grained, self-service basis over the Internet, via web applications/web services, from an off-site third-party provider who bills on a fine-grained utility computing basis. This is a general cloud available to public over Internet.

B. Private Cloud:

A private cloud is one in which the services and infrastructure are maintained on a private network. These clouds offer the greatest level of security and control, but they require the company to still purchase and maintain all the software and infrastructure, which reduces the cost savings.

C. Hybrid Cloud:

A hybrid cloud environment consisting of multiple internal and/or external providers "will be typical for most enterprises". By integrating multiple cloud services users may be able to ease the transition to public cloud services while avoiding issues such as PCI compliance.

Cloud computing is the provision of dynamically scalable and often virtualized resources as services over the internet. Users need not have knowledge of, expertise in, or control over the technology infrastructure in the cloud that supports them. Cloud computing represents a major change in how we store information and run applications. Instead of placing applications and data on an individual desktop computer, everything is hosted in the cloud, through which a collection of computers and servers accessed via the Internet.

The services offered by the cloud computing include:

1. Software as a Service(SaaS)
2. Platform as a Service(PaaS)

3. Infrastructure as a Service(IaaS)
4. Storage as a Service

Among all these services, this paper mainly focuses on Storage as a Service, and the techniques for providing security to the cloud stored data by the clients.

1) **Public auditability**: To allow TPA to verify the correctness of the cloud data on demand without retrieving a copy of the whole data or introducing additional online burden to the cloud users.
2) **Storage correctness**: To ensure that there exists no cheating cloud server that can pass the TPA's audit without indeed storing users' data intact.

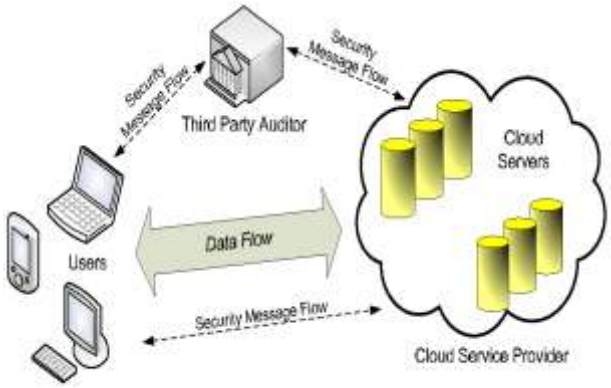


Figure: Cloud Data Storage Architecture

3) **Batch auditing**: To enable TPA with secure and efficient auditing capability to cope with multiple auditing delegations from possibly large number of different users simultaneously.
4) **Lightweight**: To allow TPA to perform auditing with minimum communication and computation overhead.

As users no longer physically possess the storage of their data, traditional cryptographic primitives for the purpose of data security protection cannot be directly adopted. In particular, simply downloading all the data for its integrity verification is not a practical solution due to the expensiveness in I/O and transmission cost across the network. Besides, it is often insufficient to detect the data corruption only when accessing the data, as it does not give users correctness assurance for those unaccessed data and might be too late to recover the data loss or damage. Considering the large size of the outsourced data and the user's constrained resource capability, the tasks of auditing the data correctness in a cloud environment can be formidable and expensive for the cloud users. Moreover, the overhead of using cloud storage should be minimized as much as possible, such that user does not need to perform too many operations to use the data.

To fully ensure the data integrity and save the cloud users' computation resources as well as online burden, it is of critical importance to enable public auditing service for cloud data storage, so that users may resort to an independent third party auditor (TPA) to audit the outsourced data when needed. To address these problems, this work utilizes the technique of public key based homomorphic linear authenticator (HLA), which enables TPA to perform the auditing without demanding the local copy of data and thus drastically reduces the communication and computation overhead as compared to the straightforward data auditing approaches. By integrating the HLA with random masking, our protocol guarantees that the TPA could not learn any knowledge about the data content stored in the cloud server during the efficient auditing process. The aggregation and algebraic properties of the authenticator further benefit this design for the batch auditing.

II. PROBLEM STATEMENT

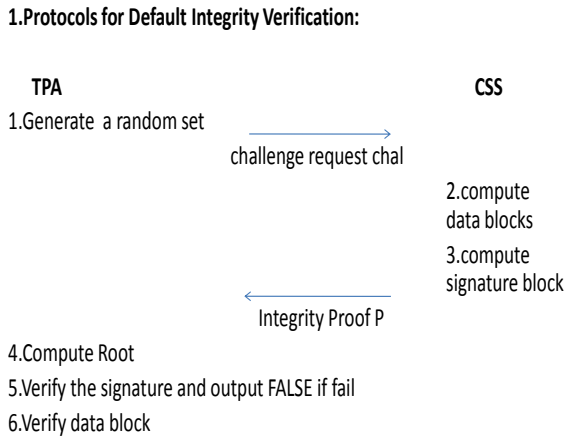
In the Existing systems, the notion of public auditability[2] has been proposed in the context of ensuring remotely stored data integrity[3] under different system and security models. Public auditability allows an external party, in addition to the user himself, to verify the correctness of remotely stored data. However, most of these schemes do not consider the privacy protection of users' data against external auditors. Indeed, they may potentially reveal user's data to auditors. This severe drawback greatly affects the security of these protocols in cloud computing. From the perspective of protecting data privacy, the users, who own the data and rely on TPA[4] just for the storage security of their data, do not want this auditing process introducing new vulnerabilities of unauthorized information leakage toward their data security.

To securely introduce an effective third party auditor (TPA), the following two fundamental requirements have to be met: 1) TPA should be able to efficiently audit the cloud data storage without demanding the local copy of data, and introduce no additional on-line burden to the cloud user; 2) The third party auditing process should bring in no new vulnerabilities towards user data privacy**.**

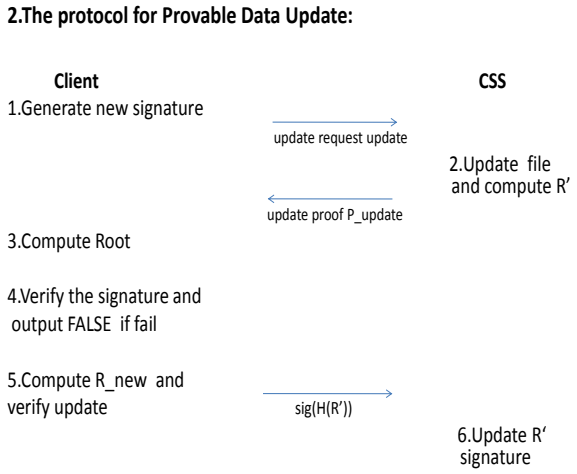
III. PROPOSED SYSTEM:

In this paper, we utilize the public key based homomorphic authenticator[5] and uniquely integrate it with random mask technique to achieve a privacy-preserving public auditing system for cloud data storage security while keeping all above requirements in mind. It includes protocols for data integrity verification and data update.

3.1 Protocols for default Integrity Verification



3.2 Protocol for Data Update



IV. HOMOMORPHICLINEAR AUTHENTICATOR:

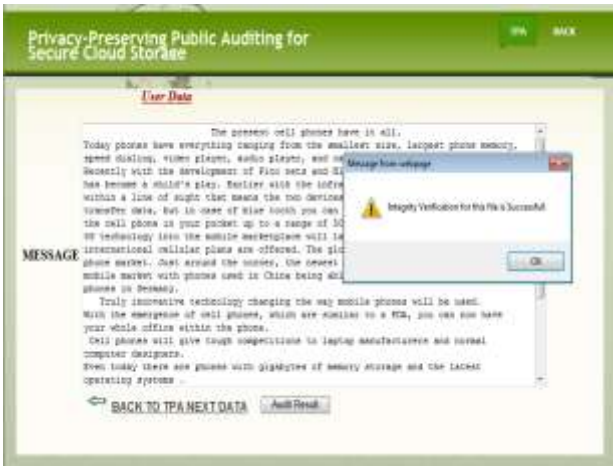
It is a way to assure an auditor that a linear combination of data blocks is correctly computed by verifying only the aggregated authenticator. It is a form of encryption which allows specific types of computations to be carried out on cipher text and generate an encrypted result which when decrypted, matches the result of operations performed on the plain text.

HLA-based Solution:

The HLA technique is used to effectively support public auditability without having to retrieve the data blocks themselves. HLAs, like MACs, are unforgivable verification metadata that authenticate the integrity of a data block. The difference is that HLAs can be aggregated. It is possible to compute an aggregated HLA which authenticates a linear combination of the individual data blocks.

At a high level, HLA-based proof of storage system works as follows. The user still authenticates each element of $F = (m1, \cdots, m_n)$ by a set of HLA. The cloud server stores $\{F, \varphi\}$. The TPA verifies the cloud storage by sending a random set of challenge $\{k\}$. (More precisely, F and $\{k\}$ are all vectors, so $\{k\}$ is an ordered set or $\{k, k_n\}$ should be sent). The cloud server then returns $\mu = Pi$ to mi and an aggregated authenticator, that is supposed to authenticate μ. Though allowing efficient data auditing and consuming only constant bandwidth, the direct adoption of these HLA-based techniques is still not suitable for our purposes. This is because the linear combination of blocks, $\mu = Pk$ to mk, may potentially reveal user data information to TPA, and violates the privacy preserving guarantee. Specifically, if an enough number of the linear combinations of the same blocks are collected, the TPA can simply derive the user's data content by solving a system of linear equations.

SCREENSHOTS:



Integrity Failure:

V. CONCLUSION:

The data storage security in Cloud Computing is discussed in this paper. This security mechanism used public auditability and homomorphism linear authenticator. The information about the data content stored on cloud server was not known by TPA. This effective and efficient auditing process not only removes the heavy load of users of the cloud. But also the user's fear of their diminish outsourced leakage of data. In this the TPA may use multiple audit sessions from various users for their outstanding data files. In a right and better way we extended our privacy-preserving public auditing protocol into multiple auditing tasks.

REFERENCES

[1] P. Mell and T. Grance, "Draft NIST working definition of cloud computing," Referenced on June. 3rd, 2009 Online at http://csrc.nist.gov/groups/SNS/cloud-computing/index. html, 2009.

[2] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in *Proc. of ESORICS'09*. Saint Malo, France: Springer-Verlag, 2009, pp.355–370.

[3] C. Wang, Q. Wang, K. Ren, and W. Lou, "Ensuring data storage security in cloud computing," in *Proc. of IWQoS'09*, Charleston,SouthCarolina,USA,2009.

[4] Shingare Vidya Marshal "Secure Audit Service by Using TPA for Data Integrity in Cloud System" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-3,Issue-4,September.

[5] Privacy-Preserving Public Auditing for Secure Cloud Storage Cong Wang, *Student Member, IEEE*, Sherman S.-M.Chow,Qian Wang,Student Member,IEEE,and Wenjing Lou,Member,IEEE.

[6] A. van Raan, "Scientometrics: State-of-the-art," Scientometrics,vol. 38, no. 1, pp. 208–218, 1996.

[7] S. Schwarze, S. Voß, G. Zhou, and G. Zhou, "Scientometric analysis of container terminals and ports literature and interaction with publications on distribution networks," in Proc. 3rd Int. Conf. Comput.Logistics, 2012, pp. 33–52.

[8] D. Straub, "The value of scientometric studies: An introduction to a debate on IS as a reference discipline," J. Assoc. Inform. Syst., vol. 7, no. 5, pp. 241–246, 2006.

[9] A. Serenko and N. Bontis, "Meta-review of knowledge management and intellectual capital literature: Citation impact and research productivity rankings," Knowl. Process Manage., vol. 11,no. 3, pp. 185–198, 2004.

[10] W. Hood and C. Wilson, "The literature of bibliometrics, scientometrics, and informetrics," Scientometrics, vol. 52, no. 2, pp. 291– 314, 2001.

[11] L. Leydesdorff, "Indicators of structural change in the dynamics of science: Entropy statistics of the SCI journal citation reports," Scientometrics, vol. 53, no. 1, pp. 131–159, 2002.

[12] S. Voß and X. Zhao, "Some steps towards a scientometric analysis of publications in machine translation," in Proc. IASTED Int. Conf. Artif. Intell. Appl., 2005, pp. 651–655.

[13] L. Leydesdorff and T. Schank, "Dynamic animations of journal maps: Indicators of structural changes and interdisciplinary developments," J. Am. Soc. Inform. Sci. Technol., vol. 59, no. 11, pp. 1810–1818, 2008.

[14] K. Sivakumaren, S. Swaminathan, and G. Karthikeyan, "Growth and development of publications on cloud computing: A scientometric study," Int. J. Inform. Library Soc., vol. 1, no. 1, pp. 37–43, 2012.

[15] Q. Bai and W.-h. Dong, "Scientometric analysis on the papers of cloud computing," Sci-Tech Inform. Develop. Econ., vol. 5, no. 1, pp. 6–8, 2011.

[16] T. Wang and G. Huang, "Research progress of cloud security from 2008 to 2011 in China," Inform. Sci., no. 1, pp. 153–160, 2013.

[17] I. Foster, Y. Zhao, I. Raicu, and S. Lu, "Cloud computing and grid computing 360-degree compared," in Proc. Grid Comput. Environ. Workshop, 2008, pp. 1–10.

[18] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," Future Gener. Comput. Syst., vol. 25, no. 6, pp. 599–616, 2009.

[19] M.Armbrust,A. Fox, R. Griffith,A.D. Joseph, R. Katz,A. Konwinski, G. Lee, D. Patterson, A. Rabkin, and I. Stoica, "A view of cloud computing," Commun.ACM, vol. 53, no. 4, pp. 50–58, 2010.

[20] C. W. Holsapple, L. E. Johnson, H. Manakyan, and J. Tanner, "Business computing research journals: A normalized citation analysis," J. Manage. Inform. Syst., vol. 11, no. 1, pp. 131–140, 1994.

[21] G. S. Howard, D. A. Cole, and S. E. Maxwell, "Research productivity in psychology based on publication in the journals of the American psychological association.," Amer. Psychol., vol. 42, no. 11, pp. 975–986, 1987.

[22] R. K. Merton, "The Matthew effect in science," Science, vol. 159, no. 3810, pp. 56–63, 1968.