International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-4, Special Issue-4, June 2016 E-ISSN: 2347-2693

A Review On Cloud Data Security With MHT Based Integrity Authentication Framework

Sajjan R.S.[1*], Vijay R. Ghorpade [2] and Arkas B.D.[3]

[1]Department of Computer Science and Engineering, University of Solapur, Solapur
[2]Department of Computer Science & Engineering, University of Kolhapur, Kolhapur
[3]Department of Computer Science and Engineering, University of Solapur, Solapur

Available online at: www.ijcseonline.org

Received: May/26/2016 Revised: Jun/02/2016 Accepted: Jun/12/2016 Published: Jun/30/2016

***Abstract*—** Cloud computing that concern the latest trend in application development for Internet services, relying on clouds of servers to manage multiple tasks that used by individual machines. So without the Internet Cloud is not for the use. The developers initiates special services, such as email, calendars, and word processing, and host them entirely online, also processing of different applications proposed by the cloud of special servers with cloud computing. Cloud storage for storing and sharing data across multiple users become more popular. Cloud computing is innovative trend that supposed to do data storing at data centres. But the limitations of cloud data security don't prescribe cloud better. To improve data reliability and availability, cloud service providers mature strategies such as storing multiple replicas along with original datasets. Public data auditing schemes enable the users to verify their outsourced data storage without retrieving the whole dataset. Due to the efficiency and security problems the system communication overhead (bandwidth) didn't maintains data integrity, and public auditing. The on-demand service provision with utilization of limited resources of client system benefits the client. However, data outsourcing paradigm in cloud is one of the biggest security concerns. Frequent integrity checking is needed to keep an eye on data. The proposed scheme makes use of Merkle Hash Tree (MHT) and AES algorithm to ensure data integrity at the unsecured server.

Keywords-Cloud, Merkle Hash Tree, Thired Party Auditor, Public Auditing

I. INTRODUCTION

Cloud is most intensive research topic that predicts the Internet services. For the management of huge upcoming data and to provide infinite computing resources on demand, the cloud is better. Cloud computing has the combination of multiple existing technologies such as web computing, parallel and distributed computing, grid computing, utility computing, virtualization etc. Cloud storage enables authentic features for customers with benefits, ranging from cost saving and simplified convenience, to mobility opportunities and scalable service. Utilize and manage the storage of user data on the cloud storage is the great feature which attracting much more customers. The cloud servers mostly utilized to relieve clients from the intensity of storage management and maintenance. Cloud computing constructs and allows us to access the applications that actually reside at remote location.

Basically cloud has the three special different definitions of cloud computing that conclude services such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Cloud computing prove advantages for organizations parsing to centralize the management of software and data storage, by guarantying of reliability and security for their users. Mostly, many efforts are covered at the commercialization of the cloud such as Amazon's in the IEEE [2].

*Corresponding Author:
Mr. Arkas B.D.
e-mail: arkasbapuraje@gmail.com , Tel.: +91-9096880864

II. RELATED WORK

Within short time people are expecting for utilizing data explosion, reading, and writing. One of the specialities utilizes cloud storage with an algorithm engine that find the exact way for data storage. The system provides a secure and performs data storage on the public cloud for use of number of users. Cloud storage has become popular by business users due to its vigorous benefits, proceeding lower cost and better resource utilization. But Cloud data storage has provided significant benefit by allowing users to store massage amount of data on demand in cost effective manner. Cloud system performs data storage on public cloud for use of number of users. Hence data security is important functional method for cloud data storage. Data integrity and authentication enhances the completeness, correctness & freshness of data [3].

Cloud is elastic computing which identify the user authentication of particular file and the system. As compared to traditional systems, scalability and elasticity are key

advantages of cloud. The efficiency has great importance in supporting dynamic data. Security and privacy protection on dynamic data has been studied especially in the past [4], [5]. In this paper, we will focus on authentic data updates, which are important because these updates exist in many cloud applications such as business transactions and online social networks (e.g. Amazon). For reliability, privacy-preserving or efficient processing and public auditing purposes Cloud users may also need to split big datasets into smaller datasets and store them in different physical servers. Integrity verification for outsourced data storage has attracted extensive research interest. Dynamic updating is prevented due to the introduction of sentinel nodes.

A. *Merkle Hash Tree (MHT)*

Merkle Hash tree is tree in which every non leaf node is acquired with the hash of the labels of its children nodes. Hash trees based on binary trees concept which are useful because they allow efficient and secure verification of the contents of larger data structures, those encrypted form of data would split into batches and those batch files are stored in cloud with the help of root hash code. The root node has the top hash key stored in local database of the owner. With the DSP, Authenticated Data Structures is a technique in which some kind of authentication data is stored on the. On the client's query, a DSP returns the queried data along with some extra authentication data that is then used by the client to verify the authenticity of returned data. In this system the main block is divided into smaller MHTs and the root hashes of these sub-trees are signed.

In Merkle proposed the use of binary trees to authenticate a large number of public keys with a single value, namely the root of the tree. That is how the definition of a Merkle tree comes into use. It is a complete binary tree with a *k*- bit value associated to each node such that the interior node value is a hash function of the main hash tree.

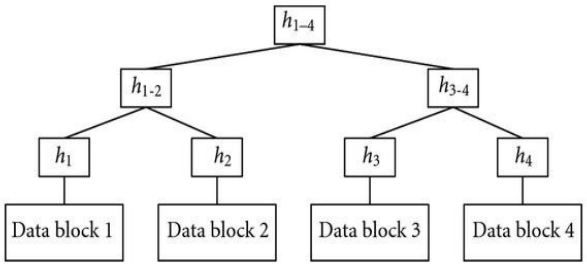

Fig1. Merkle Hash Tree structure

3.1.1. *Key generation*
In this phase we calculate the root value of the tree, and then put the first authentication path and some upcoming node values.
3.1.2. *Output*
That phase consists of *N* rounds, one for each leaf. Every round the current leaf value is output, together with the authentication path for this leaf { }*i Auth* . Tree is updated in order to prepare it for the next round.
3.1.3. *Verification*
That is the validation and verification of leaf values in a Merkle tree.

B. *Third Party Auditor (TPA)*

The TPA performs auditing on behalf of the Client. The introduction of the TPA reduces the overhead of the client. The client no longer needs to verify the integrity of the data at the server on its own. Third Party Auditor does the inspection. That considered in two categories: private audit ability and public audit ability. The private audit ability can predict higher scheme efficiency, public audit ability allows anyone, not just the client, to challenge the cloud server for the correctness of data storage while keeping none private information. To remove such lacuna of management of data of the data owner, it will audit the data of client. Achieving economies of scale for Cloud Computing that becomes eliminates the involvement of the client by auditing that whether his data stored in the cloud are indeed intact [6].

In the authorization of TPA the server enable the TPA for processing the data blocks on behalf of MHT structure. So initially the connectivity should be covered by TPA by signature scheme.

C. *Ranked Merkle Hash Tree (RMHT)*

A Merkle Hash Tree becomes authenticated data structure discussed for efficient verification of data updates by ranks, known as RMHT. On the behalf to the update algorithm, every non-leaf node will constantly have 2 child nodes. Information contained in one node N in an RMHT T is represented as {H, r_N} where H is a hash value and r_N is the rank of this node. T is constructed as follows. For a leaf node LN based on a message m_i, we have H =h (m_i), $r_{LN} = si$; A parent node of N1 = {H1, r_{N1}} and N2 = {H2, r_{N2}} is constructed as NP = {h (H1‖H2), (r_{N1} + r_{N2})} where ‖ is a concatenation operator.

D. *Verifiable Data Updating*

Typically, in the verification the client will be able to detect any fault caused by accident or dishonest behaviours in the update. In the result maintenance of verifiable update process in both our basic scheme and the modification, *Cloud Storage Server (CSS)*, i.e., R_0 cannot match the Rnew calculated by the client.

E. *Multiple Replica*

It is the efficient way of emerging in the data integrity by updating each replica one by one. The Multi-Replica scheme

is verification scheme that manipulate the effectiveness and efficiency also measure the performance of the system.

F. Cryptography

Cryptography is a method of storing and transmitting data in a particular form so that only those for whom it is intended can read and process it. The art or science encompassing the principles and methods of transforming an intelligible message into one that is unintelligible, and then retransforming that message back to its original form I wondered how information in a computer was kept safe and my friend, a computer expert, explained to me how cryptography worked. The overconfident security expert bragged endlessly about his "superior" cryptography skills, yet the sophisticated hacker group broke the code in short time.

- **Confidentiality** (the information cannot be understood by anyone for whom it was unintended)
- **Integrity** (the information cannot be altered in storage or transit between sender and intended receiver without the alteration being detected)
- **Non-repudiation** (the creator/sender of the information cannot deny at a later stage his or her intentions in the creation or transmission of the information)
- **Authentication** (the sender and receiver can confirm each other's identity and the origin/destination)

G. Hash functions

Mathematically, a hash function (or hash algorithm) is a condition of turning data into a number suitable to be handled by a computer. It enables a small digital "fingerprint" from any kind of data. The term hash apparently comes by way of analogy with its special meaning in the physical world, to "chop and mix". The function promotes or transposes the data to create that "fingerprint", usually called hash value.

H. Setup

The setup phase is used to generate security keys like private key and public key by invoking KeyGen () function. In the pre-processing, it makes use of homomorphic authenticators and Meta data. That needs to two arguments namely file and key that responsible to audit data being flown for verification of integrity of data.

The file system divides the file data into multiple blocks, Hash code is computed for each block of data by MHT structure. Especially hash code of two blocks is merged then this merged key is merged with other key made up of two merged keys. The process continues until all leaf nodes are found in the Merkle hash tree. After the processing of path the root element sent to cloud server for the data storage on cloud data server.

I. Data Integrity Verification

The TPA initiate for challenge server for block level data verification at regular intervals by sending file name and block randomly. The data integrity verification does with third party auditor. The ability of public auditing and data dynamics for remote data integrity are checked cloud computing. The system construction is conceptually continued to compute these two important goals while efficiency being also measured. To address efficient data dynamics, the system improves the existing proofs of storage models by manipulating the classic Merkle Hash Tree methodology for block tag authentication and integrity.

J. Data Modification and Data Insertion

Cloud users required to specify modifications to their online data frequently that called as data dynamics. Typically some of the existing systems, the proposed system support data dynamics. The system of cloud has to be cover Data modification and data insertion is the two important methodologies.

III. PERFORMANCE EVALUATIONS

Public data auditing schemes allow users to verify their outsourced data storage without having to retrieve the whole dataset. However, existing data auditing techniques suffers from efficiency and security problems. This section provides proposed solution by introducing a new scheme for cloud storage security. The proposed system supports both data dynamics and public auditing.

ACKNOWLEDGMENT

The authors are grateful to Ms. Sajjan R.S. Assistant professor Department of Computer Science and Engineering, Vidhya Vikas Pratisthan Institute of Engineering and Technology (VVPIET), Solapur for time to time guidance.

REFERENCES

[1] William stalling, A Handbook on "Cryptography & Network Security, Principles & Practices" Pearson Publisher, 5th edition-2011.

[2] C. Erway, A. C. Papamanthou, and R. Tamassia, "Dynamic provable data possession" in Proc.16th ACM Conf.Comput.Commun. Security, Chicago, USA, 2009, pp. 213–222.

[3] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling public audit ability and data dynamics for storage security in cloud computing," IEEE Trans. Parallel Distrib. Syst., vol. 22, no. 5,pp. 847–859, May 2011.

[4] C. Liu, J. Chen, L. T. Yang, X. Zhang, C. Yang, R. Ranjan, and K. Ramamohanarao, "Authorized public auditing of dynamic big

data storage on cloud with efficient verifiable fine-grained updates," IEEE Trans. Parallel Distrib. Syst. 2014.
[5] Chang Liu, Rajiv Ranjan, Chi Yang, Xuyun Zhang, Lizhe Wang, Senior Member " MuR-DPA: Top-Down Levelled Multi-Replica Merkle Hash Tree Based Secure Public Auditing for Dynamic Big Data Storage on Cloud " IEEE Transactions On Computers, Vol. 64, No. 9, September 2015.
[6] Yubin Xia, Yutao Liu, Haibing Guan, Yunji Chen, Tianshi Chen, Binyu Zang, Haibo Chen, "Secure Outsourcing of Virtual Appliance" IEEE Transactions on Cloud Computing DOI. 2015
[7] Christina Delimitrou and Christos Kozyrakis Security Implications of Data Mining in Cloud Scheduling", IEEE Computer Architecture Letters 2015.
[8] Jin Li, Xiaofeng Chen, Xinyi Huang, Shaohua Tang and Yang Xiang *Senior Member,* IEEE and Mohammad Mehedi Hassan *Member, IEEE* and Abdulhameed Alelaiwi *Member,* IEEE "Secure Distributed Deduplication Systems with Improved Reliability" IEEE Transactions on Computers 2015.
[9] Sumitra Samal, Barkha Agrawal, Richa Parija "A Study on Different Symmetric Key Cryptographic Algorithm" Vol. 8 No.4 Page no (1-5) Apr 2016.
[9] IEEE Explore Abstract, http://ieeexplore.ieee.org March 11 2016.
[10] Amazon Web Services, https://aws.amazon.com/, Feb 10 2016.
[11 Data security in cloud computing, https://www.youtube.com/ March 20 2016.
[12] Open Stack Open Source Cloud Software, http://openstack.org/ Feb 15 2016.

Authors Profile

Ms Sajjan R.S pursued Bachelor of Engineering in Computer Science and Engineering from Shivaji University, Kolhapur, India in the year 1999. She has pursued Master of Technology in Computer Science from PDA college of Engineering, Gulbarga, Karnatak, India in the year 2008. She is currently pursuing Ph.D. in Computer Science and Engineering from Shivaji University, Kolhapur, Maharashtra, India since 2014. She has published more than 20 research papers in reputed international journals. Her main research work focuses on Cloud Computing, Cloud Data Security, BigData, Hadoop, IoT, High performance computation cryptography. She has 15 years of teaching experience and 4 years of Research experience.

Dr. Vijay R. Ghorpade has completed his PhD in Computer Science & Engineering in 2008. He is presently working as Principal/Professor at the D. Y. Patil College of Engineering & Technology, Kolhapur, Maharashtra, India. He is guiding 8 PhD students and several PG students. He has published papers in various national and international journals. His area of interest is network security, and ad-hoc networks.

Mr. Bapu Arkas pursued Bachelor of Engineering in Computer Science and Engineering from Solapur University, Maharashtra, India in the year 2014. He is currently pursuing Master of Engineering in Computer Science and Engineering from Solapur University, Maharashtra, India since 2014. His main research work focuses on Cloud Computing, Cloud Data Security.