Data Integrity And Ensuring Computation Using MHT In Cloud Computing

Sukhaveerji Ghate$^{1,}$ Bohthala Prabhanjan Yadav$^{2,}$ S Tharun Reddy$^3$, N Swathi$^4$, Purusotham Endla$^5$

$^{1,2,4}$Sumathi Reddy Institute of Technology for Women, Warangal, India. $^3$SR Engineering College, Warangal, India.

sukhaveer.ghate222@gmail.com

Abstract: Basically, the architecture model of cloud computing provides the storage capacity as a service over the internet. The main intent of cloud computing is to provide services of security to the users and owners of data. Several security issues are appeared due to appropriate control of data. Hence in this paper the data integrity and ensuring computation using MHT in cloud computing is developed. In this first the data owner will select the initial parameters and choose the random function. The parameters are selected based on the generation of encrypted data blocks (EDB’s). After choosing function random key is computed. Now the Third Party Auditor (TPA) will initialize the data for verification. All this data is stored in the cloud server. This cloud server will follow the cloud service provider rules. Hence data which is obtained from server will computes and generate the signatures using MHT. After generation the path of the node will be identified. At last the obtained data is tested.

Keywords: Third Party Auditor (TPA), Encrypted Data Blocks (EDB’s), Merkle Hash Tree (MHT), Cloud computing.

1. Introduction

Distributed computing is an appropriated engineering model that brings together remote server assets on an adaptable stage so as to give on-request processing assets and administrations. Because of clients' prerequisites, organization models of distributed computing can vary enormously. Four sending models have been recognized, for example, Private Cloud, Community Cloud, Public Cloud, and Hybrid Cloud [1]. Moreover, contingent upon the quantity of items that are engaged with distributed computing, three design models have been perceived.

The main model incorporates the proprietor of the information and cloud specialist organization. Proprietors of information, clients or customers, and cloud specialist organizations distinguish the subsequent model [2]. At long last, a few proprietors of information, clients or customers, and cloud specialist organizations are objects in the last model. For the most part cloud specialist has three kinds of administrations they are, Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) [3].

Since distributed computing offers a mass information stockpiling, it must offer secure types of assistance also. Indeed, cloud security instrument is compelling if right protective executions come to picture. Suitable cloud security design ought to recognize the issues that will be created by security the
executives. Security is the distinctive confirmation of an affiliation's focal points (tallying information assets) trailed by the documentation, headway, and use of game plans and methodologies for using these advantages. Security the executives decides these issues with security controls which decide any disadvantages in the framework and attempt to reduce the impact of an assault. There are different kinds of security controls, for example, Deterrent Control, Preventative Controls, Corrective Control, and Detective Control [4].

Then again, the trust of distributed computing is predominantly because of the way that numerous applications and information are moving into cloud stages [6]. However, an absence of security is the significant issue for information proprietors who keep their information on cloud suppliers' servers. The significant security perspectives in distributed computing are secrecy, respectability, confirmation, approval, non-disavowal and accessibility.

The cloud servers provide optimal solutions by dealing with different capacities. This is introduced by the Feng Wang. These optimal solutions are mainly obtained in the locations of cloud platforms. By mitigating the location of the solution which was written, globalization is impacted. By using extent stream systems, capacities are migrated and extended. Generally, the overall system is deployed by using the cloud service providers and live streaming providers.

The data networking based on virtual machine is introduced by the Ruitao Xie. This will migrate the cloud data center in effective way to get exact result. By using the service providers, virtual machines are specifically migrated. By using IP addresses the physical machines are routed. Different machines are also used to migrate the services provided by the cloud platform. By using load balancing algorithm, the optimization of migration in virtual machine is done [5].

The investigation of the problem that is occurred in virtual machine is introduced by the Jiaqiang Liu. Because of this there will be reduction in the total migration time. Here the both computation and bandwidth constraints are optimized [7]. Here multiple virtual machines are migrated together by using constraint resources. The migration task will accomplish the minimum time that is taken while processing the virtual machine task.

Multiple servers based on virtual machine concept are introduced by the Yang Wang, and Menglan. This concept will host the services based on the form of replica. A parallel dynamic programming is introduced to get the formation in effective way [8-12]. Various types of request patterns also introduced in this concept to get the results effectively. The service request will be used to gratify the system results in particular format.

2. Security Model For Assuring Integrity
There are three different ways in which the security of the stored data can be compromised.

i. Sometimes, CSP can be an un-trusted, self-centered and possibly malicious. CSP may access the rarely accessed data from the cloud storage, but misrepresenting and charging the customer for the entire data.

ii. CSP may be unknown about the misuse of the users’ data by the malicious users.

iii. CSP may not announce the Byzantine failure of the storage servers in order to gain the economic benefits. Specifically, we consider four different types of attacks depending on the role of adversary.

2.1 Internal attacks:
These attacks are being introduced by the un-trusted cloud service providers or the internal employees of the organization. They are intentionally stealing the users’ data by leaking the user’s information to the outsiders.
2.2 External attacks:
These attacks are originating from the external users outside the cloud. These malicious users are capable of compromising the cloud service providers by accessing the users’ data in order to get the economic benefits.

2.3 Replace attack:
The cloud server may disposed which is based on the metadata.. consequently verifying the tested unique information metadata may be replaced based on attack.

2.4 Replay attack:
The cloud server may create the evidence from the confirmation produced for the past test, instead of recovering the tested information block in reliable way.

3. Ensuring Data Integrity
The data integrity assurance system ensures the following properties for providing the secure storage.

3.1 Public verifiability:
This scheme permits anyone, not just the data owner for checking the integrity of data stored in the cloud. Here, a Third Party Auditor (TPA) is allowed to verify integrity, thereby supporting public verifiability and private verifiability.

3.2 Privacy of data:
Even though the verification process is done by the TPA, privacy of data is assured by not leaking any information to the TPA. It is due to the fact that, the TPA is allowed to perform auditing only on the encrypted text, where the encryption keys are maintained by the data owner.

3.3 Block less verification:
For this property to be held, no whole data block should be retrieved by the TPA during the process of verification. This scheme allows the TPA to retrieve only two characters (MDB[i,j] and EDB[i,j]) at the specified position, thereby not retrieving the full data block from the cloud server. This property guarantees the block less verification of the system.

3.4 Low computation:
The amount of computation done by the data owner and the TPA is very lesser in the task of verification of the integrity of stored data in cloud server.

Low storage overhead at TPA: In this plan, the TPA or the information proprietor needs to recollect just the random function (RF) and the Secret key (Sk), in this manner limiting the measure of capacity at the TPA. In this way, the capacity overhead of TPA and information proprietor is less in the proposed model.

3.5 Security against fault:
Since the trustworthiness of calculations done by the CSP is confirmed by the information proprietor, the possibility of salami assaults is lesser.

3.6 Probability of detection of data corruption (Pd):
In this probability detection model, the issue of information debasement has been related to statistical model. The probability detection (Pd) had been researched dependent on various sorts of information debasements which are done by the un-confided in server. This information defilement can be ordered into information cancellation, information inclusion, information adjustment and information
attaching. Here, the likelihood of information inclusion, information erasure and information affixing defilements has been perceived with the most elevated system which requires a base number of challenges.

4. Data Integrity And Ensuring Computation In Cloud Computing

The below figure (1) shows the flow chart of proposed system. Integrity is a way of preserving the consistency of the stored data in cloud server and ensuring the originality of the data stored in the cloud server[13-17]. It means that the data can be modified only by authorized individuals, thereby increasing the guarantee, assurance and trustworthiness of the cloud service providers. In this first the data owner will select the initial parameters and choose the random function. The parameters are selected based on the generation of encrypted data blocks (EDB’s). After choosing function random key is computed. Now the Third Party Auditor (TPA) will initialize the data for verification. All this data is stored in the cloud server. This cloud server will follow the cloud service provider rules. Hence data which is obtained from server will computes and generate the signatures using MHT. After generation the path of the node will be identified. At last the obtained data is tested.

**Fig. 1: Data Integrity And Ensuring Computation In Cloud Computing**

i. The data file is split into nb data blocks by the data owner.

ii. All the data blocks have been encrypted using the 2-Keys symmetric encryption algorithm by the encryption service, thereby producing the encrypted data blocks EDBs.

iii. Then the EDBs are given as input to the cloud service provider.

iv. The Data owner applies some random function RF and the secret key Sk, received from the data owner over the encrypted data blocks EDBs and generate metadata blocks MDBs.

v. These Data owner are attached at the end of the encrypted data blocks EDBs by the cloud service provider.

vi. The combination of EDBs and data owner are stored together as a single data file in the cloud server by the cloud service provider.

Third Party Auditor (TPA): It refers to an organization or an individual who is having the capability to verify the integrity of the stored data in cloud server. The role of the auditor falls into two categories.
4.1. Private auditability:
In this the information proprietor is allowed to check the uprightness of the information record living
on the server.

4.2. Public auditability:
It implies that anybody, including the TPA is allowed to check the information honesty.
In this, the integrity of the re-appropriated information will be in the remote cloud server which has
been guaranteed. It has been finished by executing a strategy that gets a proof of information
ownership by producing metadata of the information in the cloud. This confirmation checks that the
information put away in the remote cloud server are not altered by unapproved clients, along these
lines initiate the information respectability.
In this MHT cloud computing plays major role to detect the path while providing the services. MHT
will generate the signatures based on the computations performed. This will mainly depend on the
servers. Hence from results it can observe that it gives effective result.

5. Results
The below figure (2) shows the bar graph for Comparison of cloud computing and MHT cloud
computing. From this it can observe that efficiency is increased and cost of this is very low.

![Fig. 2: Comparison Of Efficiency And Cost](image)

The below figure (3) shows the comparison of storage size capability for cloud computing and
MHT cloud computing. Hence it can observe the size of storage capability is very high. Hence this
system is very effective.

![Fig. 3: Comparison Of Storage Capability](image)

The below figure (4) shows the comparison of time for data generation for both cloud computing
and MHT cloud computing. Hence less time is required to generate the data compared with cloud
computing.
6. Conclusion
Hence in this paper the design and deployment of integrity assurance model that ensures the integrity of data stored in the cloud server and correctness of computations done by the cloud service provider using MHT is presented. Here the proof was obtained that the data stored in the cloud are not modified by the provider or a third-party individual, thereby ensuring the integrity of data by generating encrypted data blocks stored in cloud server. Based on the security analysis, it is demonstrated that the data protected with this security approach are relieved from the internal, external and replace attacks. From results it is observed that it gives high efficiency, occupies less storage and takes less time.

7. References
[1] Premlata Singh, Sushil Kr. Saroj, “ A Secure Data Dynamics and Public Auditing Scheme for Cloud Storage”, 978-1-7281-5197-7/20/$31.00 ©2020 IEEE
[2] Filipe Apolinário, Miguel L. Pardal, Miguel Correia, “S-Audit: Efficient Data Integrity Verification for Cloud Storage”, 2324-9013/18/31.00 ©2018 IEEE
[3] Yihui Dong, Le Sun, Dengzhi Liu, “A Survey on Data Integrity Checking in Cloud”, 978-1-5386-5059-2/18/$31.00 ©2018 IEEE
[4] Yindong Chen, Liping Li, Ziran Chen, “An Approach to Verifying Data Integrity for Cloud Storage”, 0-7695-6341-4/17/31.00 ©2017 IEEE
[5] Yong Yu, Man Ho Au, “Identity-Based Remote Data Integrity Checking With Perfect Data Privacy Preserving for Cloud Storage”, 1556-6013 © 2016 IEEE
[6] T.Subha and Dr.S.Jayashri, “Efficient Privacy Preserving Integrity Checking Model for Cloud Data Storage Security”, 978-1-5090-5888-4/16/$31.00@2016 IEEE
[7] TAN Shuang, TAN Lin, LI Xiaoling, JIA Yan, “An Efficient Method for Checking the Integrity of Data in the Cloud”,China Communications, September 2014
[8] Mr. Chandrashekhar, S. Pawar, Mr. Pankaj, R. Patil, Mr. Sujithkumar, V. Chaudhari, “Providing Security and Integrity for Data Stored In Cloud Storage”, 978-1-4799-3834-6/14/$31.00©2014 IEEE
[9] Huifeng Wang, Zhanhui Li, Xiaonan Zhao, Chanying Qi,Qinlu He,Jian Sun, “A scheme to ensure data security of cloud storage”, 978-1-4799-0530-0/13/$31.00 ©2013 IEEE
[10] P.Varalakshmi, Hamsavardhini Deventhiran, “Integrity Checking for Cloud Environment Using Encryption Algorithm”,978-1-4673-1601-9/12/$31.00 ©2012 IEEE
[11] Huaqun Wang, “Proxy Provable Data Possession in Public Clouds”, 1939-1374/12/$31.00 © 2012 IEEE
[12] Wenjun Luo, Guojing Bai, “Ensuring the Data Integrity in Cloud Data Storage”, 978-1-61284-204-2/11/$26.00 ©2011 IEEE
[13] Bikku T Rao NS and Akepogu AR 2019 A novel multi-class ensemble model based on feature selection using Hadoop framework for classifying imbalanced biomedical data International Journal of Business Intelligence and Data Mining 14(43832) 25-39 10.1504/IJBIDM.2019.096801
[14] Jayanthi N, Babu BV and Rao N.S. 2017 Survey on clinical prediction models for diabetes prediction Journal of Big Data 4(1) - 10.1186/s40537-017-0082-7
Mounika, S. & Sampath Kumar, T. 2019, "Client –requirement fulfil QOS in multi server International Journal of Advanced Science and Technology vol 28 no 7 pp 44- for max beneficial in cloud computing" 49
[15] Rajasri I, Guptha AVSSKS and Rao YVD 2016 Generation of Egts: Hamming Number
Approach Procedia Engineering 144 537-542 10.1016/j.proeng.2016.05.039

[16] Mahender K, Anil Kumar T and Ramesh KS 2017 SER and BER performance analysis of digital modulation schemes over multipath fading channels Journal of Advanced Research in Dynamical and Control Systems 9(2) 287-291

[17] Sudarshan E, Naik K.S, Kumar P.P 2020 Parallel approach for backward coding of wavelet trees with CUDA. ARPN Journal of Engineering and Applied Sciences 15(9), pp.1094-1100