International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT2390221

Data Integrity Verification in Cloud Computing

Rajesh Sada, Shatendra Dubey
Information Technology Department, NRI Institute of Information Science and Technology, Bhopal, Madhya Pradesh, India

ARTICLEINFO

Article History:

Accepted: 15 March 2023
Published: 05 April 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
208-213

ABSTRACT

High security is one of the major obstacles for opening up the new era of the long dreamed vision of Cloud computing as a utility. In today's trend of cloud computing, all the sensitive applications and data are moved towards cloud infrastructure and data center that run on virtual computing resources in the form of virtual machine. These attributes poses many security challenges such as accessibility vulnerabilities, virtualization vulnerabilities and web application vulnerabilities. We primarily aim to achieve better data integrity verification technique for Data Storage as a Service (Daas) in cloud computing and help users to recognize the threats associated with their uses.

Keywords: Third Party Auditing, Cross check , Paillier Cryptosystem

I. INTRODUCTION

Cloud computing is defined as services and applications that are enforced on a distributed network using virtualized resources and accessed by common networking standards and Internet protocols. It is distinguished by the traditional system in this manner that resources are virtual and limitless and im- plementation details of the physical systems on which softwareruns are abstracted from the user.

In cloud, the complexity of security is greatly increased in comparison with traditional system. The reason for this is that data is stored and operated in multi-tenant systems which are distributed over a wider area and are shared by unrelated users. In addition, maintenance of security audit logs may be difficult or impossible for a user that has limited resources.

The extensive use of virtualization in implementing cloud infrastructure brings unique security concerns for customers or tenants of a public cloud service. Virtualization alters the relationship between the OS and underlying hardware - be it computing, storage or even networking. This introduces an additional layer virtualization that itself must be properly configured, managed and secured. Specific concerns include the potential to compromise the virtualization software, or ”hypervisor”.

The cloud provider must ensure to cloud user that their infrastructure is secure, their data and applications are protected and it must devote proper security measures and resources to maintain privacy preservation and data integrity. However, it is possible that cloud provider may delete or sell some non operational data for its greed or profit that is not used for a long time. It is also possible that an adversary may exploit this data by performing various attacks.

Data centers are responsible for storing the data in cloud en- vironment.There are a number of security concerns associated with facility of data storage as a service of cloud computing. New security challenges introduced by storing data in the cloud are following.

1. **Data Integrity:** When a data is on a cloud, anyone from any location can access those data's from the cloud. Cloud does not differentiate between a sensitive data from a common data thus enabling anyone to access those sensitive data. Tampering the sensitive data causes the data integrity issue. Thus there is a lack of data integrity in cloud computing.
2. **Data Theft or loss:** The cloud servers are distrusted in terms of both security and reliability, which means that data may be lost or modified maliciously or accidentally. Administration errors may cause data loss (e.g., backup and restore, data migration, and changing memberships in Point to Point systems). Additionally, adversaries may initiate attacks by taking advantage of data owners' loss of control over their own data.
3. **Privacy issues:** The cloud Vendor must make sure that the Customer Personal information is well secured from other operators. As most of the servers are external, the vendor should make sure who is accessing the data and who is maintaining the server thus enabling the vendor to protect the customer's personal information.

4. **Infected Application:** Vendor should have the complete access to the server for monitoring and maintenance, thus preventing any malicious user from uploading any infected application onto the Cloud which will severely affect the customer.

Properties	**Sebe et al [1]**	**Wang et al [2]**	**Wang et al [4][3]**	**HAIL [5]**	**Hao et al [6]**
Primitives	Asymmetric-key cryptography (RSA Modules)	Merkle Hash Tree, BilinearMap	Bilinear Map, MAC, Homomorphic Authenticator	Homomorphic Verifiabl eTags	MAC, Integrity protectedECC
Type of guarantee	Probabilistic	Probabilistic	Probabilistic	Probabilistic	Deterministic
Public Verifiability	No	Yes	Yes	Yes	Yes
With the Help of TPA	No	Yes	Yes	Yes	Yes
Data dynamics	Yes	Yes	No	Yes	Yes
Privacy preserving	-	No	Yes	Yes	Yes
Support for Sampling	No	Yes	Yes	Yes	Yes
Probability of Detection	$[1\text{-}(1-p)^{c}]$	$[1\text{-}(1-p)^{c}]$	$[1\text{-}(1-p)^{c*s}]$	$[1\text{-}(1-p)^{c*s}]$	$[1\text{-}(1-p)^{c*s}]$

1. The Probabilistic guarantee of data integrity is achieved by using the probabilistic checking method. It is called probabilistic because the blocks arerandomly selected, the detection probability will be high if the server deletes a fraction of all the blocks.
2. The deterministic guarantee of data integrity is achieved by checking the integrity of all data blocks.
3. A third party auditor has certain special expertise and technical capabilities, which the clients do not have.
4. n is the block number, c is the sampling block number , s is the numbers of sectors in blocks , p and P_k is the probability of block corruption in a cloudserver and k-th cloud server in a multi-cloud P = P_k.

5. **Loss of physical control:** Cloud customers have their data and program outsourced to cloud servers. As a result, owners lose direct control on the data sets and programs. Loss of physical control

means that customers are unable to resist certain attacks and accidents. For example, data or software may be altered, lost, or even deleted; in addition, it is difficult and impractical to ensure data/computation integrity and confidentiality with traditional methods.

6. **Data Location:** In cloud environment data location is transparent from customer. The customer doesn't know where his own data's are located and the Vendor does not reveal where all the data's are stored. The data won't even be in the same country of the customer, it might be located anywhere in the world. It might raise SLA and Legal issue.
7. **Cross-VM attack via Side Channels:** Cross-VM attack exploits the nature of multi-tenancy, which enables that VMs belonging to different customers may co-reside on the same physical machine. We concentrate on the data integrity verification which is one of the biggest concerns with cloud data storage at untrusted servers because it may be possible that cloud user or /and cloud provider may be malicious. It is also an interesting problem that how cloud users and cloud providers have trusted to each other for storing the data and how privacy of the cloud users must be maintained.

One solution of this problem is to perform encryption and decryption operations but it involves with computational and operational overheads. Another solution of this problem is to perform data auditing that is a periodic event to evaluate security, data integrity, privacy preservation and computationalaccuracy.

II. LITERATURE SURVEY ON DATA AUDITING

In cloud scenario, the users might have limited computation capability, network bandwidth , battery power and communi- cation resources. So, they are not capable to perform auditing. Thus, on the behalf of cloud customer, a third party may be responsible for this data auditing task.

The schemes ([1], [2], [4], [3], [5], [6]) assigns auditing work to only a single TPA (Third Party Auditor).Table 1 shows the comparative analysis of these data auditing schemes that has single TPA.

Demerits of Data Auditing Techniques With Single TPA

- None of these scheme support TPA for cross check for the computation accuracy and data integrity verification.
- In these schemes, TPA cannot able to handle SLA and Legal issues for data possession and prone to single-point failure.
- For these schemes, error localization is very difficult to find.All the above schemes only provide binary re- sults about the storage status for identifying misbehaving server(s).
- There is a tradeoff between Data dynamics, privacy preservation and public verifiability in these schemes.

III. PROPOSED DATA AUDITING TECHNIQUE

We propose a Distributed Multiple Third Party Data Au- diting Technique. Unique features of this scheme are the following:

1) Multiple Third Party Auditors will be able to check verifiability of computation and other operation that are performed at the end of Cloud Service Provider.
2) Multiple users can simultaneously perform various dy- namic operations like update, append and delete on the same data.
3) Multiple Third Party Auditors have shared the huge load responsibilities of single TPA by load balancing.
4) Our goal is to achieve batch auditing where multiple delegated auditing tasks from different

users can be performed simultaneously by the TPA.

5) We look for homomorphic encryption with paillier cryp-tosystem.

Figure 1, presents our proposed model. In this figure, We divide Proposed Model into three Groups:

1. **Cloud Users:** Any end user may be interpreted as a cloud user. Cloud user is not capable to perform computation intensive tasks like data integrity and privacy preserving au- diting because we assume that these cloud users have limited resources. For example: Mobile customer, PDA, iphone and ipad users etc.
2. **Multiple Third Party Auditors:** In this group, TPA is an authorized and authentic entity that is responsible for data integrity verification and privacy preservation. We consider multiple TPA to achieve load balancing and batch auditing. One TPA can cross check the computation and other operation that was performed by the other TPA.
3. **Cloud Service Provider:** In this group, we assume that cloud service provider has established enough infrastructure resources like physical data centers, servers etc. to provide data storage as a service for cloud customers. These resources may be distributed across the world.

Protocol Design for Proposed Scheme is described in the following subsections.

A. Storing Mechanism

1. **Request:** Cloud user requested to the TPA group for storing the data on cloud service provider.
2. **Send:** TPA group generates key by KEY-GENERATION algorithm. Then TPA group encrypt the data content by its public key using Encryption Algorithm and send to cloud service provider.
3. **Store:** Cloud service provider decrypts these received en- crypted data contents by its own private key using Decryption algorithm and store on its own data storage centers.

B. Perform Dynamic Data Operations

1. **Request for Dynamic Data Operations over Data:** Cloud user sends requests to the TPA group for Dynamic Data Operations over the data. For this purpose, data blocks information send to TPA group.
2. **Challenge:** TPA send challenge request to cloud service provider for performing computation operation on the data sample block level of physical data centers.
3. **Response:** Cloud service provider chooses necessary data blocks from the whole data blocks , generates key by KEY- GENERATION algorithm and responded to TPA group with encrypted data content using Encryption Algorithm.
4. **Dynamic Data Operations:** TPA can apply homomorphic property of paillier cryptosys- tem [7],[8] to perform computation task like append and update operation on responded data blocks. TPA is performing this task by without retrieving any data blocks.
5. **Result Update**: The result of computation operation would be updated at the end of cloud service provider by sending a simple update request.

C. Cross Check Mechanism

1. **Request for Cross Check Computation over Data:** Cloud user sends requests to the TPA group for cross check computation over the data. For this purpose, data blocks information send to TPA group.
2. **Challenge:** TPA send challenge request to cloud service provider for performing computation operation on the data sample block level of physical data centers.
3. **Response:** Cloud service provider chooses necessary data blocks from the whole data blocks, generates key by KEY- GENERATION algorithm and responded to group TPA with encrypted data content using Encryption Algorithm.
4. **Computation for Cross Check Part-1:** TPA group performs decryption operation with its private key

on the encrypted data content using Decryption Algorithm. TPA group performs computation over these data content and the output of computation stored for verification say Result-1.

5. **Computation for Cross Check Part-2:** TPA group performs computation operation on encrypted data blocks that responded in Step 3 with the homomorphic property of our scheme and then output of the computation operation is de- crypted by TPA. We called the output of decryption operation Result-2.

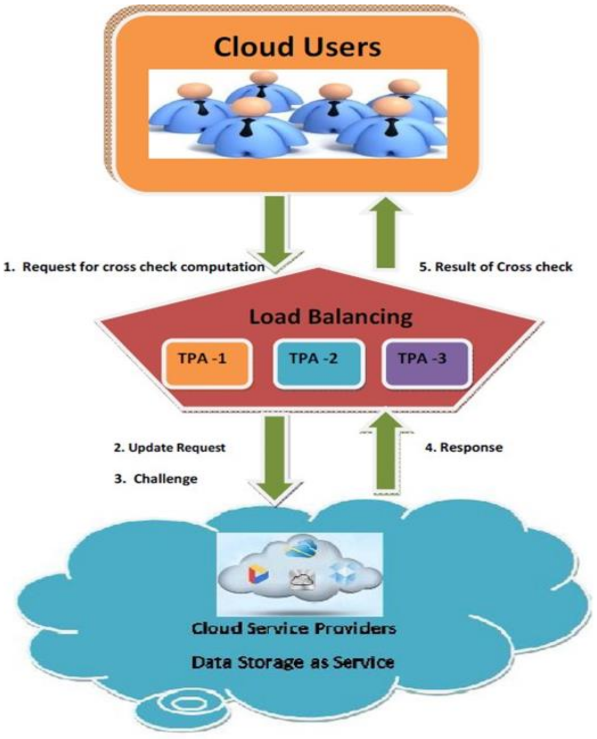

Fig. 1. Proposed Model

6. **Cross Check and Verify:** TPA matches this Result-2 with Result-1. TPA replies to cloud user the result of this crosscheck.

IV. CRYPTOGRAPHIC TOOL : PAILLIER CRYPTOSYSTEM

We used a variant of paillier cryptography system in which Group Z_N and Z_N^* are utilized such that $Z_N \times Z_N^*$ is isomorphic to $Z_{N^2}^*$.

A. Key-Generation Algorithm

1) An entity chooses two distinct odd prime numbers p andq of the same length.

2) Calculate N = pq and Euler's totient function on N is

$\varphi(N) = [(p \quad 1)(q \quad 1)]$.

3) Assure that

a) $gcd(N, \varphi(N)) = 1$.

b) For any integer $a > 0$, we have $(1 + N)^a = (1 + aN) \bmod N^2$. As a consequence, the order of (1+N) in $Z_{N^2}^*$ is N. That is, $(1 + N)^N = 1 \bmod N^2$ and $(1 + N)^a\ 1 \bmod N^2$ for any $1 < a < N$.

4) Selects a random $r \in Z_N^*$ such that gcd(L ($r^N \bmod N^2$),
N) = 1, where $L(x) = (x \quad 1)/N$.

5) The public key of our system is (N) and private key is(N, φ(N)).

A. Encryption Algorithm

Let $m \in Z_N$ be a plain-text to be encrypted and $r \in Z_N^*$ be a random number. With the definition of isomorphism, The cipher-text is given by function f that mapped into $Z_N \times Z_N^* \rightarrow Z_{N^2}^*$

$c = E (m \bmod N, r \bmod N) = f (m, r) = [(1+N)^m . r^N \bmod N^2]$

where $c \in Z_{N^2}^*$.

B. Decryption Algorithm

The user efficiently use its private key (N, φ(N))

1) Set $\hat{c} := [c^{\varphi(n)} \bmod N^2]$ where c is cipher-text .

2) Set $\hat{m} := (\hat{c} - 1) / N$. (Note that all this is carried out over the integers.)

3) After decryption plain-text is given by $m := [\hat{m} \quad \varphi(N)^{-1} \bmod N^2]$

C. Utilization of Homomorphic Property

Update Operation: On any $m1, m2 \in Z_N$ and $r1, r2 \in Z_N^*$

$$E (m1, r1).E(m2, r2) = E(m1 + m2, r1r2) mod N^2$$

We use this property for dynamic updating of data and for computing Addition of two plain-text without retrieving plain- texts.

Append Operation: On any $m1, m2 \in ZN$ and $r1, r2 \in Z_N^*$

$E^{m2}(m1, r1) = E(m1m2, r1^{m2}) modN^2$

We use this property for dynamic appending of data and for computing multiplication of two plain-text without retrieving plain-texts.

E. Self Binding Property:

With the self binding property, any cipher-text can be changed to another cipher-text without affecting the plain-text.

$$[E(m1, r1).r^{N}]_{2} modN^{2} = E(m1, r1r2)$$

We utilize this property for making the job of adversary very difficult to predict the plain-text.

V. REFERENCES

[1]. F. Sebe, J. Domingo-Ferrer, A. Martinez-Balleste, Y. Deswarte, and J. J. Quisquater. Efficient Remote Data Possession Checking in Critical Information Infrastructures. IEEE Trans. Knowledge and Data Eng,vol. 20, no. 8, pages. 1034-1038, Aug. 2008.

[2]. Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou. ”Enabling Public Verifi- ability and Data Dynamics for Storage Security in Cloud Computing,”. Proc. 14th European Conf. Research in Computer Security (ESORICS), IEEE, 2009.

[3]. C. Wang, S.S.-M. Chow, Q. Wang, K. Ren, and W. Lou ”Privacy- Preserving Public Auditing for Secure Cloud Storage”. In Cryptology ePrint Archive, Report 2009/579, 2009.

[4]. C. Wang, Q. Wang, K. Ren, and W. Lou ”Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing”. In Proc. IEEE INFOCOM, IEEE, 2010.

[5]. K.D. Bowers, A. Juels, and A. Oprea ”HAIL: A high-availability and integrity layer for cloud storage,” ,” Proc. 16th ACM conference On Computer and communications security,, pages 187-198, 2009.

[6]. Zhuo Hao, Sheng Zhong and Nenghai Yu. ”A Privacy-Preserving Remote Data Integrity Checking Protocol with Data Dynamics and Public Verifiability,” IEEE Transection on Knowledge and Data Engineering, VOL. 23, NO. 9, September 2011.

[7]. Pascal Paillier ”Public-Key Cryptosystems Based on Composite Degree Residuosity Classes” , Published in J. Stern, Ed., Advances in Cryptology EUROCRYPT’99, vol. 1592 of Lecture Notes in Computer Science, pp. 223-238, Springer-Verlag, 1999.

[8]. Pascal Paillier and David Pointcheval. ”Efficient Public-Key Cryp- tosystems Provably Secure Against Active Adversaries,” Advances in Cryptology Proceedings of ASIACRYPT , Springer-Verlag, LNCS 1716, pages 165-179,1999.

[9]. Dan Boneh and Hovav Shacham. ”CryptoBytes”, RSA Laboratories, Volume 5, No. 1 Winter/Spring 2002

[10].Dario Catalano, Rosario Gennaro , Nick Howgrave-Graham and Phong Q. Nguyen. ”Paillier’s cryptosystem revisited”, published in Proceeding CCS ’01 Proceedings of the 8th ACM conference on Computer and Communications Security Pages 206-214 , 2001

Cite this article as :

Rajesh Sada, Shatendra Dubey, "Data Integrity Verification in Cloud Computing", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 2, pp.208-213, March-April-2023. Available at doi : https://doi.org/10.32628/CSEIT2390221

Journal URL : https://ijsrcseit.com/CSEIT2390221