Volume 8, No. 5, May-June 2017
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info

A Survey Paper on Secure Auditing for Data Storing in Cloud by Third Party Auditor

Nikhil Kumar Singh
Department of Computer Science and Engineering
BabaSaheb BhimRao Ambedkar University, Lucknow, India

Abstract: IT companies are widely adopting cloud computing paradigm worldwide. It provide various Facilities to their users like remotely store huge data on server ,large scale computation and accessing of data to end user at any time and these facilities attract the user to store their data on the cloud. Outsourcing of data has become promising trends because it prevents the user's effort to heavy data maintenance and management. Outsourcing of data may be more sensitive and confidential that needs to get prevent from any type of alteration. To successfully maintain the integrity of data Auditing has proposed. Auditing is a critical aspect of an overall data integrity assurance plan to successfully audit the integrity of data. It provides the highest degree of assurance that no data integrity breaches have occurred.

Keywords: Cloud Service provider, Third Party Auditor (TPA), data integrity, Provable Data Possession (PDP).

I. INTRODUCTION

Cloud computing is widely embraced by various organization for data outsourcing. Cloud computing provides flexible and cost effective way to access outsourced data to end user in multiform without any geographical restriction. According to National Institute of Standards and technology (NIST), Cloud computing is a model for enabling worldwide, 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 managed with minimum effort or service provider interaction[1]. The basic concept behind cloud computing is virtualization; it provides virtual storage and computing service to the cloud clients. Virtualization is basically making available resources such as operating system, network, storage device and server so that they can be used by multiple users at the same time. In cloud computing the workload of users can be managed and make it more efficient, scalable and economical using virtualization.

Cloud model is composed of three service models**. First,** Software as a Service (SaaS) provides the capability to its users, to run their applications on cloud infrastructure. **Second**, Platform as a Service (PaaS) provides a platform to users to perform operations like develop, run, and manage applications. **Third,** Infrastructure as a service (IaaS) provides virtualized hardware support to its users so that they can save their investments over expensive local hardware requirements.

Cloud computing has four types of deployment models. **First**, Private cloud delivers its services same as public cloud but dedicate to single user or organisation. **Second**, Public cloud provides its services shared over multiple users and organizations. **Third,** Hybrid cloud is a combination of Public cloud and Private cloud as it works like Private cloud but can access more computing resources from third party to enhance its performance. **Fourth** is Community cloud, as its name suggests that its services are shared over multiple organizations belonging to same working area or we can say community.

There are various advantages of cloud storage services where data is remotely maintained, managed and backed up. We can outsource huge amount of data on cloud and can access from anywhere and at any time without worrying its maintenance and management and decrease the cost of hardware requirement .Due to eminent mobility and coherent storage and retrieval of data user are getting attracted to access the service of cloud [10]. As we know three main pillars of security is CIA (Confidentiality, integrity, availability). Confidentiality is hiding resources and information, it protects the content of the message and does not allow to access information to unauthorized users. Integrity refer to the trustworthiness of data or resources, it prevents unauthorized access and modifications of data. Availability refers to the skill to use the information or resources [11]. Data Owner may be worried about various security issues like the data might be accessed or altered in illegal way. In this paper we are focusing to maintain the integrity of outsourced data through auditing of that data.

Various remote data integrity verificationschemeshave been proposed to allow the auditor to check the integrity of data stored on the remotecloud server. There is basically two categories of Auditing schemes, Private Auditing and Public Auditing. Private auditing is an initial auditing model for checking integrity of outsourced data, In private Auditing scheme all computation that need for checking integrity is directly performed between data owner and cloud service provider.This scheme has its own advantage and disadvantage, its main advantage is that, it canpreserve privacy of data but the overload that increase on Data Owner side is not good at all and also it can happen that data owner and CSP both do not trust on each other about integrity proof results. Second type of Auditing is Public Auditing, in which integrity verificationprocess is done by TPA (Third Party Auditor). This scheme reduces the computation overhead of user because all computations are done through Third Party Auditor and integrity verification results produced by third party auditor are commonly accepted by both data owner and CSP.

A good Auditing scheme should have properties like privacy preserving, dynamic auditing, batch auditing, and confidentiality. Public auditing can be achieved through two basic concepts like MAC-Based and HLA-based. MAC based solution for public auditing, requires data knowledge so that privacy of data gets compromised where as HLA

based solution supports public auditing without retrieving the data information so it is privacy preserving [3]. When we think of public auditing approach, then we must have think about above mentioned issues. There should be a way to preserve privacy of user data that it cannot be revealed to TPA. There should be a facility to check integrity of dynamic data. Many schemes have been proposed for dynamic data auditing like [13],[14] and [8] etc. These schemes have achieved dynamic auditing through implementing techniques like Indexed Hash Table [5], Merkle Hash Tree [3] and Dynamic Hash Table [13]. Batch auditing is also a major concern of auditing scheme because it can enhance auditing performance in case of there exists many auditing requests from different users at the same time. [13],[3] have given techniques of batch auditing to enhance performance of their Audit process. Confidentiality of user data must be protected during audit phase. Initial PDP schemes were not privacy preserving. To achieve privacy preservation later, DPDP (MHT) [18], IHT-PA [17], DHT-PA [13] schemes have been proposed.

Public auditing allows a third party in addition to the users themselves, to check the integrity of outsourced data. We cannot fully trust on the external party as it may be honest but curious to see the data. So we can have trust on external auditing party. In this paper we have assumed that the auditor is honest over whole auditing process but it may curious to see confidential data. In addition Sometimes CSP might be dishonest. And there exists various reasons for CSP to behave unfaithfully toward the Data owner regarding their outsourced data status. For example, CSP might reclaim storage for budgetary reason by discarding data that have not been or rarely accessed, or hide data loss incident to maintain a reputation [5].In case of CSP is dishonest it may launch following attack to TPA [15]:

- **Forge attack:** The CSP may forge the data blocks and/or their tags to deceive the verifier.
- **Replace attack:**CSP can perform the replacement of corrupted data blocks and their tags with previously generated data blocks and tags so that CSP can pass the integrity check.
- **Reply attack:** The CSP may attempt to pass the verification using the proof generated from the previous ones or other former information.

System Model:

In auditing model we consider three main entities are involving they are: Data owner, Cloud service Provider and Third party auditor .The Data owner create their data and upload it on the cloud. The cloud service provider stores the data into cloud and allows accessing the data from anywhere and at any time.[24] So it is necessary to make insure that the data is same as it was uploaded by the data owner. Here is the third entities is auditor who verify the data integrity of the outsourced data for both data owner and server.[25]To verify the outsourced data, the data owner does not provide original data instead of that they give them metadata; outsourced data is almost in encrypted form. When data owner send request to TPA to check the integrity of data, the TPA send challenge to cloud service provider and regarding that challenge the CSP send the proof.[22]This way the third party auditor ensures the integrity of outsourced data.

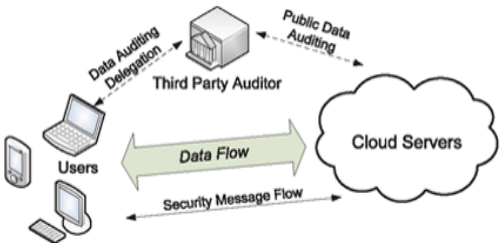

Figure 1 Security model

II. LITERATURE SURVEY

Many solutions have been proposed to check the integrity of outsourced data which can be generally divided into two categories: private auditing and public auditing. Private auditing is the beginning model for checking data integrity of outsourced data, in which data integrity checking operation can be performed between CSP and data.[1] In public auditing data verification operation is performed by TPA which reduces the overhead of Data owner. And this is a more particle way.

Proofs of Retrievability (POR) data integrity scheme is proposed by Juels et al. [4] in 2007 .this is a private auditing solution it is done by TPA cryptographic method for authenticate the integrity outsourced data stored in the cloud, without keeping a copy of the user's original files in local storage. It check the integrity of outsourced data and make sure the retrivability of data with the use of error correcting code.

PDP(Provable Data Possession) is First public auditing scheme is proposed by Ateniese et al. in 2007 ,which involve Homomorphic tags based on RSA and can remotely check the integrity of outsourced data by randomly sampling a some blocks from the file[13][15]. If differentiate with private auditing it is the first data integrity checking scheme which performed by external party not by user themselves. This scheme reduces the dispensable overhead of the user. It ensure public audit but does not have privacy preserving facility and like private audit data recovery is not supported [9].

Partially Dynamic – PDP is proposed byAtenies et al. [19] in 2008, a highly efficient and secure method for dynamic auditing based on symmetric key cryptography that not required extentencryption.it demit is it perform only limited number of audit and not support privacy preserving.

PDP(first privacy preserving PDP) is introduced by , Wang et al. [6] in 2010 presented a public auditing scheme which ensure the privacy preserving for outsourced data using Integrating the Homomorphic authenticator with random masking technique. Applied the bilinear aggregate signature to expand auditing in batch manner for multiple user, where Third party auditor can perform auditing in simultaneously manner.[2]

Cooperative PDP (CPDP) technique proposed by Zhu et al in 2012 which is scheme based on hash index hierarchy and Homomorphic verifiable scheme.[3]It Support public auditing, Privacy preserving and Batch auditing in multi cloud but it had not provision for multi user auditing.[14]

DAP (Dynamic Auditing Protocol) in 2013, Yang et al. [14] proposed further enhance auditing scheme which support dynamic auditing using the Index table scheme as data owners dynamically update their data. This paper introduced the auditing scheme for both multi user and multi

cloud to achieve batch auditing. To ensure the privacy of outsourced data they used the bilinearity property of the bilinear pairing.[5]

DPDP-MHT(dynamic provable data possession) propounded by Wang et al[18], In 2013 presented another classic public auditing scheme for dynamic auditing using Merkle Hash Tree construction for block tag authentication (MHT), to achieve efficient data dynamics. It support public auditing, Privacy preserving. Support dynamic auditing and Batch auditing in multi cloud.[12]

IHT-PA(Index Hash Table-public audit) In 2013, Zhu et al. [17] proposed further enhanced public auditing scheme based on index hash table.IN this paper auditing service formulated onrandom sampling,fragment structure and index hash table.[11]

DHT-PA (Dynamic hash table-public audit) introduced by Hui Tian et al.[22]in 2016.Dynamic hash table support public dynamic auditing and employed Homomorphic authenticator with random masking to preserve the privacy of outsourced data.

Table[1]: Comparison of existing data integrity scheme

ATA INTEGRATION SCHEME	TECHNIQUE	PROPOSED BY	YEAR	STRENGTH	WEAKNESS
POR (Proof of Retrievability)[4]	Using error correcting code	Juels et al.	2007	• Private Auditing using error code • Data recovery is possible	• Increase overhead on Data Owner. • Cannot be used in original form, pre processing is required for encoding.
PDP (provable data possession)[13]	Use Homomorphic tag based on RSA	Atenies et al.	2007	• Support public auditing	• Not Privacy preserving • No Batch auditing • Communication overhead • Data recovery is not supported
Partially Dynamic – PDP [19]	Symetric Key Cryptography	Atenies et al.	2008	• Supports Dynamic Auditing	No Privacy preserving Bounded no of Audits.
CPR (Compact Proof of Retrievability) [20]	HLA Built from secure BLS-Signature	H. Shacham, B. Waters	2008	• Improved POR scheme	• No Privacy preserving
DPDP (Dynamic PDP)[	Using ranked based authenticated skip list	Erway et al.	2009	• Dynamic data auditing • No demand of privacy preserving	• No public auditing • Not support Batch auditing • Not Privacy preserving
PDP First privacy preserving[6]	Integrating the Homomorphic authenticator with random masking	Wang et al.	2010	• Supports public auditing • Privacy preserving	
Fully Dynamic PDP [21]	Combined BLS based HLA with MHT	Wang et al.	2011	• Supports Dynamic auditing	• Not Privacy preserving
CPDP (corporative provable possession)[7]	Hash Index Hierarchy	Zhu et al.	2012	• Support public auditing • Privacy preserving • Batch auditing in multi cloud	• It does not support dynamic audit • Does not support auditing for multiuser

DAP[14]	Index table	Kan Yang et al.	2013	• Support public auditing • Privacy preserving • Support dynamic auditing • Batch auditing in multi cloud	High Computation cost
DPDP-MHT[18]	Based on Markle hash tree	Wang et al.	2013	• Support public auditing • Privacy preserving • Support dynamic auditing	• Heavy computation cost of the TPA • Large communication overhead
IHT-PA (Index hash table-public audit)[17]	Index Hash table	Zhu et al.	2013	• Support public auditing • Privacy preserving • Support dynamic auditing	Batch auditing is not mentioned
DHT-PA (Dynamic hash table-public audit)[22]	Dynamic Hash table	Hui Tian et al.	2016	• Support public auditing • Privacy preserving • Support dynamic auditing • Batch auditing in multi cloud	

III. PROPOSED SYSTEM

In our proposed system we have concentrated on time consumption acquired in setup phase at the data owner side of whole audit system. As we know that privacy of outsourced data must not be reveled at TPA side at any cost. So we need a powerful setup mechanism so that privacy preservation can be achieved. Setup phase for large files may be tedious and time consuming task for data owner.Existing audit system works like explained ahead. At first stage called setup phase, Data is divided into blocks of equal size. Tags or Authenticators are generated for each data block at next stage. Further TAGs or Authenticators are stored on cloud server along with their blocks. When data owner needs to check integrity of his outsourced data then he will raise a request to TPA. TPA then forwards another request to CSP for data proof according to request of data owner. It may be for whole data blocks or for some set of data blocks of a file requested for audit operation. CSP then generates data proof and send back the proof to TPA. TPA then verifies the data proof sent by CSP and sends the result of audit as true/false to the data owner.In the whole above described process, we can analyze that data owner faces a tedious and time consuming operation at setup phase to achieve privacy preservation.To overcome this problem we can use multithreading concept to reduce time consumption of setup phase . Prior setup phases were based on serial execution of generating data tags and authenticators. As we know about multiprocessing capability of CPU, Multithreading approach will boost the process of generating TAGs, Authenticators for each data block because threads will execute simultaneously. This approach will certainly reduce the time consumption of setup phase. None of prior implementationshave used this proposed concept. We have considered in this paper that we can improve setup phase throughput by using multithreading technique that we would say “Multithreaded Setup Phase” and it may be a great achievement ahead in cloud auditing system in future.

IV. CONCLUSION

In this paper we have analyzed different types of PDP schemes on the basis of Privacy preservation, dynamic auditing, batch auditing, communication and computation overhead.In some PDP schemesthey have achieved privacy preservation by adopting some complex data setup operations like generating tags or authenticators for each data block. None of the single previous schemes has mentioned the solution about overhead of setup phase at data owner side. We have proposed a solution for above mentioned problem through implementation of multithreading technique in setup phase. This approach will reduce the time consumption of data setup process at data owner side.

V. REFERENCES

1. P. Melland T. Grance, "The NIST Definition of Cloud Computing, "technical report, Nat'lInst. of Standards and Technology, 2009.
2. C. Wang, Q. Wang, K. Ren and W. Lou, ''Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing, '' Proc. IEEE INFOCOM, pp. 1-9, 2010.
3. C. Wang, S. M. Chow, Q. Wang, K. Ren and W. Lou, ″Privacy-Preserving Public Auditing for Secure Cloud Storage,″ IEEE Trans. on Computers, vol. 62, no. 2, pp. 362-375, 2013.
4. A. Juels and B.S. Kaliski Jr., "PoRs: Proofs of Retrievability for Large Files", Proc. ACM Conf. Computer and Communications Security (CCS '07), pp. 584-597, 2007.
5. K. Yang and X. Jia, "An Efficient and Secure Dynamic Auditing Protocol for Data Storage in Cloud Computing," IEEE Trans. on Parallel and Distributed Systems, vol. 24, no. 9, pp.1717-1726, 2013.ISSN: 2278 – 1323
6. C. Wang, Q. Wang, K. Ren and W. Lou, ''Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing,'' Proc. IEEE INFOCOM, pp. 1-9, 2010.
7. Y. Zhu, H. Hu, G. Ahn, and M. Yu, "Cooperative Provable Data Possession for Integrity Verification in Multi-Cloud Storage," IEEE Trans. Parallel and Distributed Systems, vol. 23, no. 12, pp. 2231-2244, 2012
8. Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu and S. S.Yau, "Dynamic Audit Services for Outsourced Storage in Clouds," IEEE Trans. on Services Computing, vol. 6, no. 2, pp. 227–238, 2013.
9. International Journal of Computer Applications (0975 – 8887) Volume 122 – No.2, July 2015 27 A Survey on Data Integrity Techniques in Cloud Computing.
10. K. Shinde, V. V. Jog, "A Survey on Integrity Checking for Outsourced Data in Cloud using TPA", International Journal of Computer Applications (0975 – 8887), International Conference on Internet of Things, Next Generation Networks and Cloud Computing.
11. J. wei Li, J. Li, D. Xie, and Z. Cai, "Secure Auditing and De-Duplicating Data in Cloud", 2386 IEEETRANSACTIONSONCOMPUTERS,VOL.65,NO. 8,AUGUST2016.
12. Sumalatha.M.R., Hemalathaa.S., Monika.R., Ahila.C., "towards secure audit services for outsourced data in cloud", 2014 International Conference on Recent Trends in Information Technology 978-1-4799-4989-2/14/$31.00 © 2014 IEEE.
13. Hui Tian, Member, IEEE, Yuxiang Chen, Chin-Chen Chang, Fellow, IEEE, "Dynamic-Hash-Table Based Public Auditing for Secure Cloud Storage", IEEE TRANSACTIONS ON SEVICE COMPUTING, MANUSCRIPT ID.
14. Ch. Mutyalanna, p. Srinivasulu, m. Kiran3, "dynamic audit service outsourcing for data integrity in clOUDS", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2 Issue 8, August 2013.
15. G. Ateniese, R.B. Johns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson and D. Song, ''Provable Data Possession at Untrusted Stores,'' Proc. 14th ACM Conf. on Comput. and Commun. Security (CCS), pp. 598-609, 2007.
16. Y. Zhu, H. Hu, G. Ahn, and M. Yu, "Cooperative Provable Data Possession for integrity Verification in Multi-Cloud Storage," IEEE Trans. Parallel and Distributed Systems, vol. 23, no. 12, pp. 2231-2244, 2012.
17. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu and S. S.Yau, " Dy-namic Audit Services for Outsourced Storage in Clouds," IEEETrans. on Services Computing, vol. 6, no. 2, pp. 227–238, 2013.
18. Q. Wang, C. Wang, K. Ren, W. Lou and J. Li. ''Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing,'' IEEE Trans. on Parallel and Distributed Systems, vol. 22, no. 5, pp. 847-859, 2011
19. G. Ateniese , R. D. Pietro, L. V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession", Proc. Int ' lConf. Security and Privacy in Comm. Networks (Secure Comm'08), pp. 1-10, 2008.
20. H. Shacham and B. Waters, "Compact Proofs of Retrievability," Proc. Int'l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (Asiacrypt), ol. 5350, pp. 90-107, Dec. 2008.
21. Q.Wang,C.Wang,K.Ren,W.Lou,andJ.Li,"EnablingPublic AuditabilityandDataDynamicsforStorageSecurityinCloud Computing," IEEETrans.ParallelandDistributedSystems, vol.22,no.5,pp.847-859,May2011.
22. H Tian, Y Chen, C Chang, H Jiang, O Huang, Y Chen, J Liu, "Dynamic-Hash-Table Based Public Auditing for Secure Cloud Storage"
23. S Lins, S Schneider, and A Sunyaev, "Trust is Good, Control is Better: Creating Secure Clouds by Continuous Auditing",IEEE TRANSACTIONS ON CLOUD COMPUTING, TCC-2015-10-0378
24. A Kushanpalli, V. S. Kumar, C. R. Yadav, "A Simulation Study of Outsourcing of Audit Service for Data Integrity in Cloud Computing", ISSN (Print) : 2319-5940 International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 11, November 2014.
25. D. N. Rewadkar, S. Y. Ghatage, "Cloud Storage System Enabling Secure Privacy Preserving Third Party Audit", International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014.
26. S. Pearson, "Toward Accountability in the Cloud", IEEE Internet Computing, vol. 15, no. 4, pp. 64–69, 2011.
27. Cloud Security Alliance,"Top Threats to Cloud Computing," http://www.cloudsecurityalliance.org, 2010.
28. C. Wang, K. Ren, W. Lou, and J. Li, Towards Publicly Auditable Secure Cloud Data Storage Services, IEEE Network Magazine, vol. 24, no. 4, pp. 19-24, July/Aug. 2010.