An examination on data integrity auditing patterns in cloud computing

Article *in* International Journal of Engineering & Technology · March 2018

DOI: 10.14419/ijet.v7i1.9.9822

2 authors, including:

Mohanaprakash T.A
Panimalar Institute of Technology
81 PUBLICATIONS **452** CITATIONS

International Journal of Engineering & Technology, 7 (1.9) (2018) 200-204

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

An examination on data integrity auditing patterns in cloud computing

T. A. Mohanaprakash [1] *, J. Andrews [2]

[1] *Associate Professor, Department of CSE, Panimalar Institute of Technology, Chennai*
[2] *Professor, Department of CSE, Sathyabama Institute of Science and Technology, Chennai*
**Corresponding author E-mail: tamohanaprakash@gmail.com*

Abstract

Cloud computing is associate inclusive new approach on however computing services square measure made and utilized. Cloud computing is associate accomplishment of assorted styles of services that has attracted several users in today's state of affairs. The foremost enticing service of cloud computing is information outsourcing, because of this the information homeowners will host any size of information on the cloud server and users will access the information from cloud server once needed. A dynamic outsourced auditing theme that cannot solely defend against any dishonest entity and collision, however conjointly support verifiable dynamic updates to outsourced information. The new epitome of information outsourcing conjointly faces the new security challenges. However, users might not totally trust the cloud service suppliers (CSPs) as a result of typically they may be dishonest. It's tough to work out whether or not the CSPs meet the customer's expectations for information security. Therefore, to with success maintain the integrity of cloud information, several auditing schemes are projected. Some existing integrity ways will solely serve for statically archived information and a few auditing techniques is used for the dynamically updated information. The analyzed numerous existing information integrity auditing schemes together with their consequences.

Keywords: *Integrity; Cloud Computing; Cloud Service Suppliers; Third Party Auditor.*

1. Introduction

Cloud computing is widely embraced by many organization and individuals because of its various dazzle advantages like huge size data storage, cumbersome computation, low price service and flexible way to access the data [1]. The basic concept behind cloud computing is virtualization. In cloud computing, virtualization means to create a virtual variation of a device or resource, such as a server, storage device, network or operating system where the structure divides the resource into required number of execution environments [2, 12].

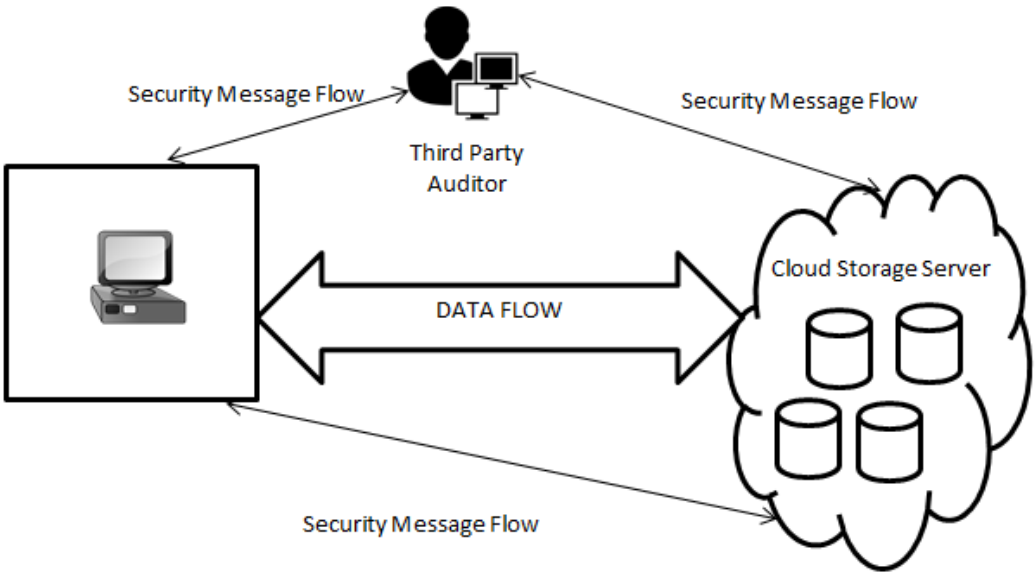

Fig. 1: Cloud Storage Architecture.

Cloud computing is a predominant service of cloud storage, which allows data owner to store their data from their local computing system to cloud. Many users store their data on cloud storage. However new protocol of data hosting service also introduces security issue. Data owner would be worry that data could be lost in the cloud [3, 14]. Therefore, the biggest concern is how to determine whether a cloud storage system and service provider meet the customer expectations for data security [4].

Therefore, it is crucial and significant to amplify efficient auditing scheme to strengthen data owners′ faith in cloud storage. Various types of auditing models have been proposed, they can be categorized into two types Private auditing model and Public Auditing Model [5]. Traditionally in Private auditing model data owner can verify the integrity of outsourced data based on the two-party storage auditing protocol. In this technique data owner should have expertise. It increases the overhead of data owner and sometimes it also happens both data owner and CSP cannot convince each other for the result [6, 13].

As public auditing is the advisable model for outsourced data verification, it additionally involves the third party to check the integrity which can provide equitable auditing result for both data owner and CSP. Data owner send metadata to TPA instead of original data. Basically, auditing model has two phases set up phase and verification [7]

2. Related work

In the contemporary year, cloud storage auditing has attracted attention to strengthen data owners' trust and confidence in cloud storage. To verify the integrity of outsourced data many protocols have been proposed with distinct techniques [8]. The first auditing related work was introduced in 2007 by Juels et al. is POR (Proof of Retrievability) scheme, which can check the correctness of data with the use of error correcting code. It is typically a private auditing model because there is no existence of any other third party [11]. In the same year, Atenies et al has introduced first public Auditing Model, PDP using Homomorphic tag based on RSA. It does not support privacy preserving of data [9].

Beside data integrity auditing there are many other significant concerns such as privacy- preserving, batch auditing, and dynamic auditing. In 2008, Atenies et al. has further proposed the scheme which supports dynamic auditing but does not preserve privacy. In 2009 Erway et al. proposed dynamic PDP scheme that does not require privacy preserving. In 2010, First privacy preserving PDP was introduced by Wang et al they presented a public auditing scheme which ensures the privacy preserving for outsourced data using integrating Homomorphic authenticator with the random masking technique. In 2012 further, a new public auditing scheme Cooperative PDP (CPDP) technique proposed by Zhu et al, which was based on hash index hierarchy and Homomorphic verifiable scheme. It Supports public auditing, Privacy preserving and Batch auditing in the multi cloud but it has no provision for multi-user auditing. Dynamic Auditing Protocol (DAP) in 2013, Yang et al. proposed further enhanced auditing schemes which supported dynamic auditing using the Index table scheme. In 2015, Identity-Based Distributed Provable Data Possession (ID-DPDP) scheme was proposed by Wang, Huaqun which used bilinear pairing in random access model. Dynamic Hash Table-Public Audit (DHT-PA) introduced by Hui Tian et al. in 2016 proposed Dynamic hash table which supported public dynamic auditing. Dynamic hash table supports public dynamic auditing and employed Homomorphic authenticator with random masking to preserve the privacy of outsourced data. They used aggregate BLS signature to arrange batch auditing [10].

3. Literature survey

S.NO	TITLE	AUTHOR	CONCEPT	YEAR	ADVANTAGE	DISADVANTAGE
1	Unequal Failure Protection Coding Technique for Distributed Cloud Storage Systems	Yupeng Hu. Yonghe Liu. Wenjia Li	A deep reinforcement learning framework that integrates a 3-D convolutional neu ral network, a deep Q-network, and a residual recurrent neural network for the efficient semantic parsing of large-scale 3-D point clouds. The proposed framework provides an end-to-end automatic processing method that maps the raw point cloud to the classification results of the given categories..	2017	Data recovery is possible	Increased data cloud preprocessing.
2	Unequal Failure Protection Coding Technique for Distributed Cloud Storage Systems	Yupeng Hu Yonghe Liu	Our erasure coding technique makes the first attempt to take advantage of the unequal failure rates across the disks/nodes to optimize the system reliability and reconstruction performance. Specifically, our proposed technique, the Unequal Failure Protection based Local Reconstruction Code (UFP-LRC) divides the data blocks into several unequal-sized groups with local parities, assigning the data blocks stored on more failureprone disks/nodes into the smaller-sized group, so as to provide unequal failure protection for each group.	2017	optimization	Not privacy preservation
3	A Survey on Metrics and Measurement Tools for Sustainable Distributed Cloud Networks	Ana Carolina Riekstin. Bastos Rodrigues	We present an extensive study of metrics, methods, and tools to support sustainable operations in distributed cloud networks, with the aim of providing an end-to-end and up-to-date scenario to support current and coming research, as well as to analyze existing gaps.	2017	Support operations in distributed	No privacy preservations
4	The Internet of Microgrids: A Cloud Based Framework for Wide-Area Networked Microgrids	Eric Joseph Harmon Utku Ozgur Mehmet Hazar Cintuglu	This paper presents a cloud-based and hybrid wireless mesh communication framework for bi-level, nested, distributed optimization of networked clusters of microgrids. The proposed optimization framework implements a diffusion-based, fully distributed algorithm on	2017	wide-area internet-based cloud	No privacy preservations

			local wireless network and a quasi-distributed approach on wide-area internet-based cloud. The lower level of the bi-level optimization implements a distributed optimal economic dispatch solution for intra-microgrid among distributed energy resources, and the upper level implements a global optimal dispatch for inter-microgrid energy exchange. To demonstrate industrial applicability of the proposed framework, the IEC 61850 interoperability protocol is adopted to achieve a certain delay performance so that the distributed optimization			
5	A Survey: Smart agriculture IoT with cloud computing	Mahammad Shareef Mekala. P. Viswanathan	Precision agriculture sensor monitoring network is used greatly to measure agri-related information like temperature, humidity, soil PH, soil nutrition levels, water level etc. so, with IoT farmers can remotely monitor their crop and equipment by phones and computers. In this paper, we surveyed some typical applications of Agriculture IoT Sensor Monitoring Network technologies using Cloud computing as the backbone. This survey is used to understand the different technologies and to build sustainable smart agriculture. Simple IoT agriculture model is addressed with a wireless network.	2017	Monitoring Network technologies using Cloud	No privacy preservations
6	Remote patient health monitoring cloud brokering services	Raafat Aburukba. Assim Sagahyroon. Mohammed Elnawawy.	Exist commercialized cloud platforms that deal with the integration of patient devices that are manufactured by the same provider. This work proposes a healthcare brokering services that enables the integration with existing cloud platforms to capture the data from the patient's devices. Moreover, the work also presents a way to model the patient's health condition to be remotely monitored and takes right decision at the right time. This work is validated by a prototype implementation.	2017	integration with existing cloud platforms to capture the data from the patient's	No privacy preservations
7	Integrity and confidentiality preservation in cloud	Kirti Dhwaj Singh. Ayushi Sharma	We provide a scheme which main-tains the integrity and confidentiality of logs as well as provides easy verification mechanism for the logs. Various encryption techniques and hash functions are used to provide integrity to the generated logs in the system. We use RESTful API to publish our logs on the web for verification in the system and also by the auditor. To avoid non-repudiation, we use digital signature generation. Our scheme ensures that even the cloud provider cannot collude with the attacker. Unlike the techniques introduced so far for storing and verifying logs, we introduce bloom sequence which ensures a low false positive rate.	2017	provide integrity to the generated logs in the system.	attacks are used to bring down different types of cloud service models
8	Dynamic replica creation strategy based on file heat and node load in hybrid cloud	YaHui Zhao. ChunLin Li LaYuan Li	Combined with the node load, the average heat and the average load are used to adjust the number of copies in this paper, which can adapt to the changes of the environment dynamically. Experiments show that with changes of file access and traffic intensity, the improved strategy is sensitive to access frequency, which can adaptively adjust the number of copies, reduce the average response time, and achieve better load balance of cluster.	2017	access and traffic intensity	No privacy preservations
9	Fake File Detection in P2P Networks by Consensus and Reputation	Paul Watters Robert Layton	Social Disc is a new kind of mass-storage characterized by unlimited space and provides the set of operations on files: creation, reading, deletion and modification. Furthermore, the attempts to improve the operation of StegHash by trade-off between memory requirements and computation time are presented: a)	dynamic verification of free addresses to the multimedia objects.2017	an application of linked list. dynamic verification of free addresses to the multimedia objects.	The features, limitations and opportunities of the design were discussed.

			an application of linked list structure; b) introduction of a The features, limitations and opportunities of the design were discussed.			
10	Metrics Visualization Technique Based on the Origins and Function Layers for OSS-Based Development	Ryosuke Ishizue Hironori Washizaki Yoshiaki Fukazawa	Software developments involving multiple organizations such as OSS (Open Source Software)-based projects tend to have numerous defects when one organization develops and another organization edits the program source code files. Developments with complex file creation, modification history (origin), and software architecture (functional layer) are increasing in OSS-based development.	2017	numerous defects when one organization develops and another organization edits the program source code files.	attacks are used to bring down different types of cloud service models
11	PPFS: A Scale-Out Distributed File System for Post-Petascale Systems	Fuyumasa Takatsu Kohei Hiraga Osamu Tatebe	The prototype implementation of PPFS with file creation optimization achieves 119,000 operations per second for file creation when using five metadata servers and 128 client processes, thereby exceeding the performance of IndexFS 2.17 times. With local access optimization, PPOSS reached its limit at a block size of 16 KiB, an improvement of 1.5 times compared to before optimization. Furthermore, this evaluation indicates that PPFS has scalability on file creation and IO performance, that is required for post-petascale systems.	2017	optimization, PPOSS reached its limit at a block size	No privacy preservations
12	File sharing system encapsulated with customized social networking and learning management system	Aratrika Sarkar Nikhil Prakash	. A root is at the topmost position in the organization with levels below him. The application enables a user to create files and share them with other users depending upon his or her position in the hierarchy. This application encapsulates several security measures integrated with the flexibility of sharing files easily with a single or multiple users without the use of `email'. This application is simple, easy to use and secure. Security of files is implemented by the use of cryptography in various file modes. In this paper, we have created an application which aids in file sharing within an organization coupled with a security system.	2017	measures integrated with the flexibility of sharing files easily with a single or multiple users without the use of `email'.	No privacy preservations
13	Time Identification of Electronic Documents	Zhang Guoyou	By studying the domestic and foreign literature, from the current research situation that only solved the problems of creation time, latest modification time and latest access time. For the problems of how to make sure the files time when its had been changed, copied, deleted, recovered for many times, how to make sure it's creation time and upload time on the web front platform before entering the server, how to make sure the files creation time when the system time is not right as well as how to set and implement the time sign management standards during the files creation process etc., which need further research. Many technical methods of electronic document's time identification is given in the paper.	2017	. Many technical methods of electronic document's time identification is given in the paper.	No privacy preservations
14	Using server-to-server communication in parallel file systems to simplify consistency and improve performance	Philip H. Carns Bradley W. Settlemyer Walter B. Ligon	In this paper we examine the use of collective communication between the storage servers to improve the scalability of file metadata operations. In particular, we apply server-to-server communication to simplify consistency checking and improve the performance of file creation, file removal, and file stat. Our results indicate that collective communication is an effective scheme for simplifying consistency checks and significantly improving the performance for several real metadata intensive workloads.		server-to-server communication to simplify consistency checking and improve the performance of file creation	No privacy preservations

15	Notice of Retraction Creating user-relationship-graph in use of flow-net and log files for computer and network accountability and forensics	Daisuke Takahashi Yang Xiao Ke Meng	In log files, not all information/events are recorded and it is thus impossible to trace the paths of secret leaking based on log files alone. In this paper, we utilize user-relationship-graphs, or social networks, to compensate for the required information. User-relationship-graphs are constructed from several flow-net data structures over a longer period so that we can avoid missing embedded threats such as hostile codes. We call this approach virtual flow-net.	2017	User-relationship-graphs are constructed from several flow-net data structures	No privacy preservations

4. Conclusion

In cloud computing, a new paradigm of data outsourcing increases new security challenges. The new paradigm requires a Third-Party Auditor to check the data integrity in cloud storage. Here, we have compared different types of auditing schemes on the basis of Privacy preservation, dynamic auditing and batch auditing along with their strength and weakness. From all these papers, it is concluded that there is need to design some optimizing techniques that can be applied to speed up the set phase at data owner side. We have proposed a multi-threading model on multi-core CPU system to generate the signature for each block it is one- time operation and occurs in setup phase at data owner side.

5. Future work

In future, we focus on enhanced & sophisticated data setup process to reduce the computation and communication overhead at data owner side. To generate authenticators, we use multi-threading framework on latest multi-core system to speed up the setup phase. We use the multi-threading model in each step of data setup phase.

References

[1] J. Meza, Q. Wu, S. Kumar, and O. Mutlu, “A Large-Scale Study of Flash Memory Failures in the Field,” in Proc. SIGMETRICS, pp. 177–190, 2015.

[2] M. Shahidehpour, Z. Li, S. Bahramirad, Z. Li and W. Tian, "Networked Microgrids: Exploring the Possibilities of the IIT-Bronzeville Grid," in IEEE Power and Energy Magazine, vol. 15, no. 4, pp. 63-71, July-Aug. 2017.

[3] M. Shahidehpour, Z. Li, S. Bahramirad, Z. Li and W. Tian, "Networked Microgrids: Exploring the Possibilities of the IIT-Bronzeville Grid," in IEEE Power and Energy Magazine, vol. 15, no. 4, pp. 63-71, July-Aug. 2017.

[4] M. H. Cintuglu; T. Youssef; O. A. Mohammed, "Development and Application of a Real-Time Testbed for Multiagent System Interoperability: A Case Study on Hierarchical Microgrid Control," in IEEE Transactions on Smart Grid, vol.PP, no.99, pp.1-1, 2015.

[5] M. Shahidehpour, Z. Li, S. Bahramirad, Z. Li and W. Tian, "Networked Microgrids: Exploring the Possibilities of the IIT-Bronzeville Grid," in IEEE Power and Energy Magazine, vol. 15, no. 4, pp. 63-71, July-Aug. 2017.

[6] R. Duan, R. Prodan, X. Li, "Multi-Objective Game Theoretic Scheduling of Bag-of-Tasks Workflows on Hybrid Clouds [J]", IEEE Trans on Cloud Computing, vol. 2, no. 1, pp. 29-42, 2014.

[7] Rajesh, M., “A Review on Excellence Analysis of Relationship Spur Advance in Wireless Ad Hoc Network." international Journal of Pure and Applied Mathematics Volume 118 No. 9, 407-412, 2018.

[8] D. Takahashi, Y. Xiao, "Retrieving Knowledge from Auditing Log Files for Computer and Network Forensics and Accountability", Security and Communication Networks, vol. 1, no. 2, pp. 147-160, Feb. 2008.

[9] A. Devulapalli and P. Wyckoff, "File creation strategies in a distributed metadata file system," Parallel and Distributed Processing Symposium, 2007. IPDPS 2007. IEEE International, pp. 1-10, 26-30 March 2007.

[10] Rajesh, M., and J. M. Gnanasekar. "Path Observation Based Physical Routing Protocol for Wireless Ad Hoc Networks." Wireless Personal Communications 97, no. 1: 1267-1289. (2017)

[11] K. Hiraga, O. Tatebe, "Design and implemantation of distirbuted metadata server using non-blocking transaction for distributed file system", IPSJ SIG Technical Reports of High Performance Computing (HPC), vol. 2012, no. 28, pp. 1-9, 2012.

[12] Wei Wang, Lei Chen, Qian Zhang, “Outsourcing high-dimensional healthcrae data to cloud with personalized privacy preservation”, Computer Networks, Volume 88, 9, Pages 136–148, 2015.

[13] K. Vijayakumar C, Arun, Continuous security assessment of cloud based applications using distributed hashing algorithm in SDLC,Cluster Computing DOI 10.1007/s10586-017-1176-x,Sept 2017.

[14] K. Vijayakumar C,Arun, Analysis and selection of risk assessment frameworks for cloud based enterprise applications”, Biomedical Research, ISSN: 0976-1683 (Electronic), January 2017.