ISSN: 2321-7782 (Online)
Volume 2, Issue 5, May 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Paper
Available online at: www.ijarcsms.com

Robust Data Integrity Mechanism for Outsourced Cloud Data

U. Ranjani[1]
II M.E. (CSE),
Shivani Engineering College,
Trichy, Tamilnadu India.

M. Nathiya[2]
B.Tech. (IT), M.E. (CSE), Assistant Professor,
Shivani Engineering College,
Trichy, Tamilnadu, India.

Abstract: ***An external auditor to audit user's outsourced data in the cloud without any modification on the data content. The data hosting service introduces new security challenge, that auditing cannot be applied to the data in the cloud directly from the data owners that can be applied through the TPA. For that to designed an auditing framework for cloud storage system, an efficient and secured auditing service with a third party authority. The user only communicates with the TPA. The auditing supports the data dynamic operations. First to support scalable and efficient public auditing in the cloud computing. Safer and Snefru technique makes the transaction of data to cloud in secure manner. In particular, that scheme achieves auditing where multiple delegated auditing tasks from different users can be performed simultaneously by the TPA. Achieving the Data Dynamics is also important. To reduce the computation speed and ram utilization of the auditor, it is desirable to combine all these responses together and do the batch verification.***

Index Terms: Storage auditing; dynamic auditing; batch auditing; privacy-preserving auditing; cloud computing.

I. INTRODUCTION

Cloud computing is an expression used to describe a variety of computing concepts that involve a large number of computers connected through a real-time communication networks such as the Internet. Cloud computing is a synonym for distributed computing over a network, and means the ability to run a program or application on many connected computers at the same time.

The phrase also more commonly refers to network-based services, which appear to be provided by real server hardware, and are in fact served up by virtual hardware, simulated by software running on one or more real machines. The popularity of the term can be attributed to its use in marketing to sell hosted services in the sense of application service provisioning that run client server software on a remote location. Cloud computing relies on sharing of resources to achieve coherence and economies of scale similar to a utility over a network. At the foundation of cloud computing is the broader concept of converged infrastructure and shared services. The cloud also focuses on maximizing the effectiveness of the shared resources. Cloud resources are usually not only shared by multiple users but as dynamically re-allocated per demand. That can work for allocating resources to users.

The most basic cloud-service model, providers of IaaS-Infrastructure as a Service offer computers physical or virtual machines and other resources. A hyper visor, such as Xen or KVM or VMware ESXESXi, runs the virtual machines as guests. Pools of hypervisors within the cloud operational support-system can support large numbers of virtual machines and the ability to scale services up and down according to customers' varying requirements. IaaS clouds often offer additional resources such as a virtual-machine disk image library, block and file-based storage, firewalls, load balancers, IP addresses, virtual local area networks VLANs, and software bundles. IaaS cloud providers supply these resources on-demand from their large pools installed in data centers. For wide-area connectivity, customers can use either the Internet or carrier clouds, dedicated virtual private networks.

II. RELATED WORK

[a] Merkle Hash Tree Construction

To improve the existing proof of storage models by manipulating the classic Merkle Hash Tree construction for block tag authentication [8]. To support efficient handling of multiple auditing tasks. The technique of bilinear aggregate signature to extend the main result into a multi-user setting, where the TPA can perform multiple auditing tasks simultaneously. Extensive security and performance analysis show that the schemes are highly efficient and provably secure.

A general formal PoR model with public verifiability for cloud data storage, in which block-less verification is achieved and equip the PoR construction with the function of supporting for fully dynamic data operations, especially to support block insertion, which is missing in most existing schemes [12]. Then also prove the security of construction and justify the performance of that scheme through concrete implementation and comparisons with the state-of-the-art. A dynamic auditing protocol may leak the data content to the auditor because it requires the server to send the linear combinations of data blocks to the auditor.

TPA may not concurrently handle multiple audit sessions from different users for their outsourced data files, in the cloud paradigm, by putting the large data files on the remote servers, the clients can be relieved of the burden of storage and computation It is critical importance for the clients to ensure the data that their being correctly stored and maintained. Hash functions and signature schemes cannot work on the outsourced data without a local copy of data.

[b] Cooperative Provable Data Possession

To support distributed cloud storage, to illustrate a representative architecture used in the cooperative PDP scheme architecture has a hierarchy structure which resembles a natural representation of file storage [10]. The hierarchical structure consists of to represent relationships among all blocks for stored resources. Homomorphic verifiable response is the key technique of CPDP because it not only reduces the communication bandwidth, but also conceals the location of outsourced data in the distributed cloud storage environment.

The PDP or PoR is such a probabilistic proof technique mainly focus on PDP issues at untrusted servers in a single cloud storage but not suitable for a multicloud environment [11]. The data possession without downloading data at untrusted stores, they are not suitable for a distributed cloud storage environment since they were not originally constructed on interactive proof system. The homomorphic responses doesn't responses from multiple clouds can be also combined into a single value on the client side [2]. For lack of, homomorphic responses clients must invoke the PDP protocol repeatedly to check the integrity of file blocks stored in multiple cloud servers.

[c] Privacy-Preserving Public Auditing

A system based on HLA, which covers many recent proof of storage systems. That will pinpoint the reason why all existing HLA based systems are not privacy-preserving [3]. The analysis of these basic schemes leads to the main result, which overcomes all these drawbacks. The linear combination of sampled blocks in the server's response is masked with randomness generated the server [3], [7]. The random masking, the TPA no longer has all the necessary information to build up a correct group of linear equations and therefore cannot derive the user's data content, no matter how many linear combinations of the same set of file blocks can be collected.

To extend their dynamic auditing scheme to be privacy preserving and support the batch auditing for multiple owners. However, due to the large number of data tags, their auditing protocols will incur a heavy storage overhead on the server. To achieved privacy up to certain extent but which increases complex key management on user side. .In cloud, data is stored in a centralized form and managing this data and providing security is a difficult task.

TPA requires knowledge of data blocks for verification, Communication & computation complexity. Limitation on data files to be audited as secret keys are fixed, after usages of all possible secret keys, the user has to download all the data to recomputed MAC & republish it on CS. TPA should maintain & update states for TPA which is very difficult and it supports only for static data not for dynamic data.

[d] Homomorphic Authenticators

The data owners can remotely store their data in the cloud to enjoy on-demand high quality applications and services from a shared pool of configurable computing resources [7]. While data outsourcing relieves the owners of the burden of local data storage and maintenance, it also eliminates their physical control of storage dependability and security, which traditionally has been expected by both enterprises and individuals with high service-level requirements.

Homomorphic Authenticators to significantly reduce the arbitrarily large communication overhead for public auditability without introducing any online burden on the data owner [2], Homomorphic authenticators are unforgeable metadata generated from individual data blocks, which can be securely aggregated in such a way to assure a verifier that a linear combination of data blocks is correctly computed by verifying only the aggregated authenticator.

The legal issues in areas such as,regulatory compliance and auditing, all of which have not been well understood. TPA can see the actual content stored on a cloud during the auditing phase. TPA itself may leak the information stored in the cloud which violate data security. That is still not enough for a publicly auditable secure cloud data storage system.

III. SYSTEM MODEL

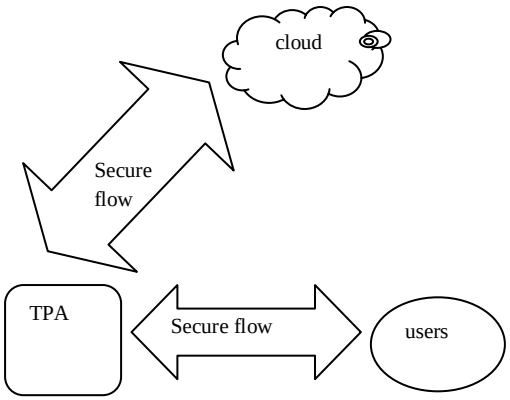


Fig.1. System model of the data storage auditing

The model gives the definition of storage auditing protocol. To designs an auditing framework for cloud storage systems and proposes a privacy-preserving and efficient storage auditing protocol. That auditing protocol ensures the data privacy by using cryptography method and convinces both cloud service providers and the data owners. And also we extend our auditing protocol to perform dynamic auditing to support the data dynamic operations. Dynamic auditing helps to ensure privacy to the dynamic data that are stored in the cloud [9]. To utilize the Strong cryptographic technique and uniquely integrate with a random mask technique to achieves a privacy-preserving public auditing system for cloud data storage security while keeping all above requirements in mind.

The main advantage of this dynamic auditing protocol is to audit and ensure security for the data that are dynamically accessed by the data owners in the cloud. It also protects the data privacy against the auditor by combining the cryptography method [9]. Data stored and accessed in the cloud can be secured and the auditor convinces the both cloud service provider and data owner [6]. This auditing scheme incurs less computation cost of the auditor by moving the computing loads of auditing from the auditor to the server, which greatly improves the auditing performance.

IV. MODULE DESCRIPTION

1. Key Generation

Key generation is the process of generating cryptography keys to security using DES technique. The Data Encryption Standard (DES) is a previously predominant symmetric-key algorithm for the encryption of electronic data. It was highly influential in the advancement of modern cryptography in the academic world. DES is a block cipher with a block size of 64 bits, so the output will be a multiple of 64 bits long. It's possible that this is stored directly in your sample hash, but unlikely, not least because that would mean that the "MD5" in your sample hash line would be meaningless.

2. Assigning Key To File

First the server map the keys to files and the TPA encrypt the files using PBEWithMD5AndDES algorithom that PBE stands for password-based encryption, generate a symmetric key using MD5 in this case from a password then use that key to encrypt something perhaps encrypt the user's password with DES. PBE with MD5 and DES password-based encryption. That algorithm entails using a password, a byte array known as salt, and an iteration count along with an MD5 message digest to produce a DES secret key; this key is then used to perform DES encryption or decryption. Primarily to encrypt the private keys, although it may be used to encrypt any arbitrary data. Corresponding key, than store the keys and in a hash table. Because accessing the data using index is less complexity. They cannot do search the whole data just we search index of the data. so that process will be very speed. Finally, the users access the data using that secret key.

3. Data Storage On Cloud Server

Cloud Computing moves the application software and databases to the large data centers, the management of the data and services may not be fully trustworthy. The unique attribute, however, poses many new security challenges which have not been well understood. In the article, to focus on cloud data storage security, Store the encrypted files in a different location on a cloud server, the requester only having that corresponding key and the requester gives that key to the TPA, then the TPA will use that key checks the data verification, but the TPA does not see the original data, only checks the validation using the signature scheme in cryptography.

4. Tpa Integrity Verification

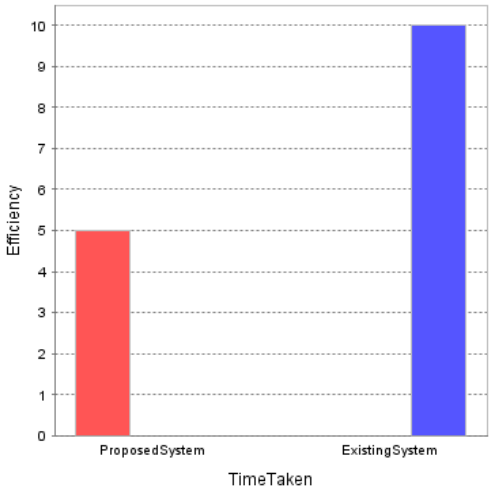
Decrypt the each and every file in a cloud server and Combine all the files. TPA Check the data size and the size will be same of original data. If any data loss occur for technical problem in a particular file, then put the corresponding encrypt file in that location, and reporting to content owner. They cannot loss security, to store the entire file in a encrypt format. Hash code is generating for the received each token from the cloud.

5. Dynamic Data Verification

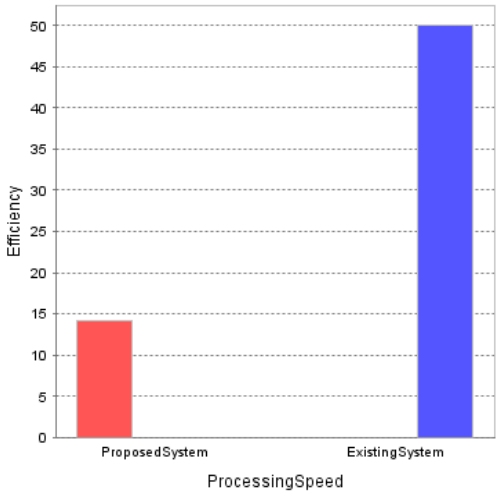
The dynamic auditing protocol that can support the dynamic operations, the data owners will dynamically update their data. The owner completes the data dynamic operations; it sends an update message to the auditor, when for updating the data that is stored on the auditor. After the confirmation auditing, the auditor sends the result to the owner for the confirmation that the owner’s data on the server and the abstraction information on the auditor are both up-to-date.

6. Batch Auditing

Multiple users auditing request is handled simultaneously. It also supports batch auditing through which efficiency is improved. It allows TPA to perform multiple auditing task simultaneously and it reduces communication and computation cost. Through that scheme can identify invalid response. System performance will be faster. In a cloud server, lot of users store their files. So each user validates their data using batch system. The auditing time will be very less. Data storage auditing is a significant service in cloud computing that helps the owners to check the data integrity on the cloud servers.



a) Time Taken



b) Processing Speed

Fig.2. Simulation Results

V. CONCLUSION

The audit protocol effectively maintains the integrity of the data that are stored in the cloud through the TPA. An external auditor to audit user's outsourced data in the cloud without any modification on the data content. The user only communicates with the TPA, so the user is always free. The batch auditing protocol can also the batch for multiple owners. Furthermore, the auditing scheme incurs less communication cost and less computation cost of the auditor by moving the computing loads of auditing from the auditor to the server, which greatly improves the auditing performance and can be applied to large-scale cloud storage systems. The powerful combination of high-assurance remote server integrity mechanism to reduce the computation speed and ram utilization of the auditor. Extensive security and performance analysis using Safer and snefru algorithm.

ACKNOWLEDGMENT

The authors would like to thank Kan Yang, Student Member, IEEE, and Xiaohua Jia, Fellow, IEEE

References

1. Kan Yang, Student Member, IEEE, and Xiaohua Jia, Fellow, IEEE "An Efficient and Secure Dynamic Auditing Protocol for Data Storage in Cloud Computing" IEEE transactions on parallel and distributed systems, vol. 24, no. 9, september 2013.
2. C. Wang, K. Ren, W. Lou, and J. Li, "Toward Publicly Auditable Secure Cloud Data Storage Services," IEEE Network, vol. 24, no. 4,pp. 19-24, July/Aug. 2010.
3. C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing," Proc. IEEE INFOCOM, pp. 525-533, 2010.
4. G.R. Goodson, J.J. Wylie, G.R. Ganger, and M.K. Reiter, "Efficient Byzantine-Tolerant Erasure-Coded Storage," Proc. Int'l Conf.Dependable Systems and Networks, pp. 135-144, 2004.
5. G. Yamamoto, S. Oda, and K. Aoki, "Fast Integrity for Large Data," Proc. ECRYPT Workshop Software Performance Enhancement for Encryption and Decryption, pp. 21-32, June 2007.
6. K. Yang and X. Jia, "Data Storage Auditing Service in Cloud Computing: Challenges, Methods and Opportunities," World Wide Web, vol. 15, no. 4, pp. 409-428, 2012.
7. K. Zeng, "Publicly Verifiable Remote Data Integrity," Proc. 10th Int'l Conf. Information and Comm. Security, L. Chen, M.D. Ryan, and G. Wang, eds., pp. 419-434, 2008.
8. 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. Parallel Distributed Systems, vol. 22, no. 5, pp. 847-859, May 2011.
9. Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S.S. Yau, "Dynamic Audit Services for Integrity Verification of Outsourced Storages in Clouds," Proc. ACM Symp. Applied Computing, W.C. Chu, W.E. Wong, M.J. Palakal, and C.-C. Hung, eds., pp. 1550-1557, 2011.
10. 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, Dec. 2012.
11. G. Ateniese, R.C. Burns, R. Curtmola, J. Herring, L. issner, Z.N.J. Peterson, and D.X. Song, "Provable Data Possession at Untrusted Stores," Proc. ACM Conf. Computer and Comm. Security, P. Ning, S.D.C. di Vimercati, and P.F. Syverson, eds., pp. 598-609, 2007.
12. A. Juels and B.S. Kaliski Jr., "Pors: Proofs of Retrievability for Large Files," Proc. ACM Conf. Computer and Comm. Security, P. Ning, S.D.C. di Vimercati, and P.F. Syverson, eds., pp. 584-597, 2007.

13. G. Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," IACR Cryptology ePrint Archive, vol. 2008, p. 114, 2008.

AUTHOR(S) PROFILE

U.RANJANI received the B.E. degree in Computer Science and Engineering from Shri Angalamman college of Engineering And Technology, Tiruchirappalli, Anna University, Chennai, Tamilnadu, India and the M.E. Computer Science and Engineering from Shivani Engineering College, Tiruchirappalli, Anna University, Chennai, Tamilnadu, India in 2012 and 2014.

