International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 04 | Apr-2018 www.irjet.net p-ISSN: 2395-0072

A Survey on Remote Data Possession Verification Protocol in Cloud Storage

Sandesh R[1], Shashi Rekha H[2]

[1]*Fourth Sem M.Tech, Department of Studies in CS&E, VTU PG Center, Mysuru*
[2]*Assistant professor, Department of Studies in CS&E, VTU PG Center, Mysuru*

Abstract – *As a leading cloud computing application, cloud storage offers scalable, flexible, and high quality services for data storage and processing. An increasing number of data owners are choosing to outsource data files to the cloud. Because cloud storage servers are not completely reliable, data owners need reliable tools to verify that their files are outsourced to remote cloud servers. To solve this crucial problem, a Remote Data Possession Verification (RDPV) protocol has been introduced. But many existing systems have vulnerabilities in data efficiency or dynamics. In this article, we provide a new efficient RDPV protocol based on the homomorphic hash function. The new system is surely secure against forgery attacks, replacing attacks by attack and copy based on a typical security model. To support data dynamics, an ORT (Operations Record Table) is introduced to track file block operations. We also provide a new optimized implementation for ORT that makes ORT access costs almost constant. In addition, we perform a complete performance analysis that shows that our system has advantages in terms of calculation and communication costs. The implementation of the prototype and the experiments show that the scheme is feasible for real applications.*

***Key Words*: Cloud computing, RDPV: Remote Data Possession Verification, ORT: Operation Record Table, Data Dynamics, Homomorphic Hash Functions.**

1. INTRODUCTION

Cloud computing appears as a new paradigm of calculation after grid computing. By managing a large number of computing resources across the Internet, it has a huge computing capacity and virtualized storage space. As a result, cloud computing is widely accepted and used in many real world applications. As a leading cloud service, the cloud service provider provides users with reliable, scalable and low-cost outsourced storage services. It provides users with a more flexible means called pay-as-you-go model for computing and on-demand storage resources. Based on this model, users can engage the necessary IT infrastructure for initial user investments that will be significantly reduced. In addition, it is convenient for them to adjust the capacity of the leased resource by changing the scale of their applications.

The cloud service provider tries to provide a promising service for data storage, saving users the investment and resource costs. However, cloud storage also involves various security issues for outsourced data. Although some security issues have been resolved, important issues related to data falsification and data loss are still present in cloud storage. On the one hand, the crash disk error or the cloud storage server (CSS) hardware failure can cause unexpected corruption of outsourced files. On the other hand, CSS is not completely reliable from the point of view of the data owner; it can actively delete or modify files for huge economic benefits. At the same time, CSS can hide the data owner's incorrect behavior and data loss incidents in order to maintain a good reputation. Therefore, it is crucial that the data owner uses an effective means of verifying the integrity of data outsourcing.

Remote Data Possession Verification Protocol (RDPV) is an effective technique for ensuring the integrity of data files stored on CSS. RDPV provides a data owner method to effectively check whether the cloud service provider is faithfully storing the original files without recovering them. In RDPV, the data owner can challenge the CSS on the integrity of the destination file. CSS can generate evidence to show that it maintains complete and incorrect data. The basic requirement is that the data owner can verify the integrity of the files without accessing the complete original file. Another desired requirement is that dynamic data operations must be supported by the protocol. In general, the data owner can add, insert, delete, or modify file blocks as needed. In addition, the complexity of the calculation and the communication overhead of the protocol must be taken into account for real applications.

2. RELATED WORK

Cloud computing is an emerging model in which IT infrastructure resources are provided as a service on the Internet. Data owners can outsource their data by remotely storing them in the cloud and enjoying high-quality on-demand services from a shared pool of configurable computing resources. However, because data owners and cloud servers are not in the same trusted domain, outsourced data can be at risk because the cloud server is no longer fully reliable. Therefore, the confidentiality, availability and integrity of the data are of crucial importance in such a scenario. The data owner encrypts the data before storing it on the cloud to ensure data privacy. The cloud should allow owners or a trusted third party to verify the integrity of their data storage without requiring a local copy of the data. Owners often replicate their data across cloud servers across multiple datacenters to provide a higher level of scalability, availability, and durability. When

data owners request the cloud service provider (CSP) to replicate data, they are charged higher storage fees by the CSP. Therefore, data owners must strongly believe that the CSP stores copies of data agreed in the SLA, and that data updates have been successfully performed on all remote-stored copies. To deal with such problems, previous multi-copy verification schemes focused on static files or incurring huge update costs in a dynamic file scenario. In this paper, we propose a Dynamic Multi-Replicable Proving Data Possession Scheme (DMR-PDP) that, while preserving data confidentiality, prevents the CSP from cheating by keeping fewer copies than paid and / or falsified. In addition, we are also expanding the schema of support for a basic file versioning system where only the difference between the original file and the updated file is propagated rather than the propagation of operations for reasons of confidentiality. DMR-PDP also supports efficient dynamic operations such as block modification, insertion and deletion on replicas on cloud servers. Through safety analysis and experimental results, we demonstrate that the proposed system is secure and works better than other recently published ideas [2].

In cloud computing, customers put large data files on the cloud storage server unreliable, how to ensure that the integrity of outsourced data becomes a big problem. To solve this problem, Hao et al. proposed a protocol that supports public verifiability and data dynamics. However, the protocol of Hao et al. suffers two disadvantages. First the protocol of Hao et al. Not sure and cannot resist the active opponent. Second, the protocol of Hao et al. Supports only fixed-size blocks as a base unit. As a result, the insertion of a short message will result in a considerable waste of storage space. In this article, we propose an enhanced remote data integrity verification protocol that can support blocks of varying size. In addition, the enhanced protocol can withstand the attack of the active adversary, and it is obvious to verify that the enhanced protocol still preserves the properties of the Hao et al. Like the public audit and the confidentiality audit [9].

While storage outsourcing and resource sharing networks have become popular, the problem of effectively proving the integrity of data stored on untrusted servers has received increased attention. In the Proving Data Possession (PDP) model, the client preprocesses the data and then sends it to an insecure server for storage, while retaining a small amount of metadata. The client then asks the server to prove that the stored data has not been falsified or deleted (without downloading actual data). However, the original PDP schema applies only to static files (or append-only) [3].

The definition framework and effective constructs for Proven Data Retention (DPDP), where it extends the PDP model to support provable updates to stored data. We use a new version of authenticated dictionaries based on ranking information. The price of dynamic updates is a change in performance from O (1) to O (log n) (or O (n ǫ log n)), for a file consisting of n blocks, while maintaining the same probability (or better, respectively) detecting misbehavior. Our experiments show that this slowdown is very low in practice (for example, proof size of 415 KB and calculation overhead of 30 ms for a 1 GB file). We also show how to apply our DPDP system to outsourced file systems and version control systems (for example, CVS).

Many storage systems rely on replication to increase the availability and durability of data on unreliable storage systems. At present, such storage systems provide no solid evidence that multiple copies of the data are actually stored. Storage servers can get along to give the impression that they store many copies of the data, when in reality they only store one copy. We are remedying this gap through the possession of multiple replicable provable data (MR-PDP): A secure schema that allows a client that stores replicas of a file in a storage system to verify through a replica protocol. Only one can be produced at the time of the challenge and the storage system uses one time the storage required to store a single replica. MR-PDP extends the previous work on proof of data ownership for a single copy of a file in a client / server storage system. Using MR-PDP to store replicas is much more computationally efficient than using a single-replica PDP schema to store unrelated separate files (for example, by encrypting each file separately before store it). Another advantage of MR-PDP is that it can generate other replicas on demand, at low cost, when some existing replicas fail [4].

The HAIL (High-Availability and Integrity Layer), a distributed cryptographic system that allows a set of servers to prove to a client that a stored file is intact and recoverable. HAIL strengthens, formalizes and streamlines the distinct approaches of cryptographic communities and distributed systems. The HAIL proofs are efficiently calculable by the servers and very compact - usually tens or hundreds of bytes, regardless of the size of the file. HAIL checks cryptographically and reactively reallocates the file shares. It is robust against an active and mobile opponent, that is to say capable of progressively corrupting all servers. We propose a strong and formal accusatory model for HAIL, as well as a rigorous analysis and choice of parameters. We show how HAIL improves the security and efficiency of existing tools, such as proof of retrievability (POR) deployed on individual servers. We also report on a prototype implementation [8].

3. PRELIMINARIS

A. System Framework

In cloud computing, cloud storage provides the user with scalable, flexible and high quality data storage and computing services. An increasing number of data owners are choosing to outsource data files to the cloud. Since cloud storage servers are not completely reliable, data owners need reliable ways to verify ownership of their outsourced files on remote cloud servers. We provide a new and efficient RDPV protocol based on the homomorphic hashing function. The new system is surely secure against falsification attacks,

replaces attacks, and replays attacks based on a typical security model. To support data dynamics, an ORT (Operation Record Table) is introduced to track file block operations. the cloud storage system consisting of two participants: CSS and data owner. CSS has powerful storage and compute capabilities, it accepts data owner requests to store externalized data files and provides access service. The data owner benefits from the CSS service and puts a large amount of CSS files without backup copies locally. As the CSS is not supposed reliable and sometimes behaves badly.

B. Remote Data Possession Checking Protocol

Remote Data Possession Verification (RDPV) is an effective technique for ensuring the integrity of data files stored on CSS. RDPV provides a method for the data owner to effectively check if the cloud service provider is faithfully storing the original files without recovering them. In RDPV, the data owner can challenge the CSS on the integrity of the target file. CSS can generate evidence to prove that it keeps the data complete and uncorrupted. The basic requirement is that the data owner can perform file integrity checking without accessing the complete original file. In addition, the protocol must withstand the malicious server attempting to verify the integrity of the data without accessing the complete and uncorrupted data. Because CSS is not supposed to be reliable and sometimes behaves badly, for example by modifying or deleting partial data files, the data owner can effectively check the integrity of the outsourced data. An RDPV schema includes the following process. They are KeyGen, TagGen, Challenge, ProofGen, Check, PrepareUpdate, ExecUpdate.

C. Homomorphic Hash Function

The RDPV scheme uses a homomorphic hash function technique, in which the hash value of the sum for two blocks is equal to the product for two hash values of the corresponding blocks. We introduce a linear table called ORT to record data operations to support data dynamics such as block modification, block insertion, and block deletion. To improve the efficiency of access to ORT, the table is reserved on the data owner side and used to record all dynamic behaviors on the file blocks. We use a doubly linked list and matrix to present an optimized implementation of ORT that reduces the cost to an almost constant level. We prove that the presented system is secure against forgery attacks, replay attacks, and replace attacks based on a typical security model.

4. CONCLUSION

In this article, we examine the problem of controlling the integrity of remote-processed data files to the remote server, and we propose an efficient and secure RDPV protocol with dynamic data. Our schema uses a homomorphic hash function to check the integrity of the files stored on the remote server and reduces the storage costs and computational costs of the data owner. We are designing a new lightweight RDPV protocol to support dynamic block operations that result in minimal computational costs by reducing the number of node moves. Thanks to our new data structure, the data owner can perform file insertion, modification or deletion operations with great efficiency. The presented schema is safe in the existing security model. The results of the experiment will indicate that our system is convenient in cloud storage

ACKNOWLEDGEMENT

The author would like to thank Dr. K Thippeswamy, Professor and chairman, Dept. of studies in computer science and engineering, VTU Regional office, Mysuru and anonymous reviewers encouragement and constructive piece of advice that of prompted us for new round of rethinking of our research, additional experiments and clearer presentation of technical content.

REFERENCES

[1]. R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic,"Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," Future Gener. Comp. Sy., vol. 25, no. 6, pp. 599 – 616, 2016.

[2]. R. Mukundan, S. Madria and M. Linderman, ''Efficient integrity verification of replicated data in cloud using homomorphic encryption,'' Distrib. Parallel Dat., vol. 32, no. 4, pp. 507-534, 2014.

[3]. C. Erway, A. Küpçü, C. Papamanthou, and R. Tamassia, "Dynamic Provable Data Possession,'' in Proc. 16th ACM Conf. on Comput. And Commun. Security (CCS), 2009, pp. 213-222.

[4]. R. Curtmola, O. Khan, R. Burns, and G. Ateniese, ''MR-PDP: Multiple-replica provable data possession,'' in Proc. 28th IEEE Conf. on Distrib. Comput. Syst. (ICDCS), 2008, pp. 411-420.

[5]. J. Li, Y. Shi and Y. Zhang, "Searchable ciphertext-policy attribute-based encryption with revocation in cloud storage," Int. J. Commun. Syst., DOI: 10.1002/dac.2942.

[6]. J.G. Han, W. Susilo, Y. Mu and J. Yan, "Privacy-Preserving Decentralized DKiesytr-iPboultiecdy SAystttreimbust,e v-Bola. s2e3d, nEon.1c1ry, pptpio. n2,1"5 0IE-2E1E62 ,T 2r0an1s2a ctions on Parallel and

[7]. Z. J. Fu, X. M. Sun, Q. Liu, L. Zhou, and J. G. Shu, "Achieving efficient cloud search services: multi-keyword ranked search over encrypted cloud data supporting parallel computing," IEICE Transactions on Communications, vol. E98-B, no. 1, pp.190-200, 2015.

[8]. K. D. Bowers, A. Juels, and A. Oprea, "Hail: A high-availability and integrity layer for cloud storage," in

Proc. 16th ACM Conf. on Comput. and Commun. Security (CCS), 2009, pp. 187–198.

[9]. E. Zhou and Z. Li, ''An improved remote data possession checking protocol in cloud storage,'' in Proc. 14th Int'l Conf. on Algs. and Archs. for Parall Proc. (ICA3PP), 2014, pp. 611-617

[10]. T.J. Ren, J. Shen,J. G. Wang, "Mutual verifiable provable data auditing in public cloud storage,"

BIOGRAPHIES

Sandesh R presently pursuing his M.Tech degree in department of studies in CSE at Visvesvaraya Technological University, PG Center, Mysuru 570029. He has completed B.E in CSE branch at Vidyavardhaka College of Engineering, Mysuru, Karnataka in the year 2016. His M.Tech project area is on Cloud Computing. This paper is a survey paper of his M.Tech project.

Shashi Rekha H presently working as Asst. Professor in DOS in CS & E , VTU-PG Center , Mysuru since 2013. Her qualification is B.E, M.tech, (Ph.D). She has 11 years of teaching experience. She is currently pursuing research in the area of Big Data analytics. Her research interests are Image classification, Pattern recognition, Data Mining in E- healthcare. She has presented many papers in various journals and few conferences. She is a member of CSI, Research Gate Forum.