

AN EFFICIENT AND FLEXIBLE OF IDENTITY-BASED DISTRIBUTED PROVABLE DATA POSSESSION IN MULTI-CLOUD STORAGE

¹E.Fathima, ²R.Velvizhi, ³K.Sivaraman

^{1,2,3} Assistant Professor, Dept. Of CSE, BIST, BIHER, Bharath University

¹fathima.cse@bharathuniv.ac.in

Abstract: Remote data integrity checking is of crucial importance in cloud storage. It can make the clients verify whether their outsourced data is kept intact without downloading the whole data. In some application scenarios, the clients have to store their data on multi cloud servers. At the same time, the integrity checking protocol must be efficient in order to save the verifier's cost. From the two points, we propose a novel remote data integrity checking model: ID-DPDP (identity-based distributed provable data possession) in multi cloud storage. The formal system model and security model are given. Based on the bilinear pairings, a concrete ID-DPDP protocol is designed. The proposed ID-DPDP protocol is provably secure under the hardness assumption of the standard CDH (computational Diffie-Hellman) problem. In addition to the structural advantage of elimination of certificate management, our ID-DPDP protocol is also efficient and flexible. Based on the client's authorization, the proposed ID-DPDP protocol can realize private verification, delegated verification, and public verification.

1. Introduction

Cloud computing has become an important theme in the computer field. Essentially, it takes the information processing as a service, such as storage, computing. It relieves of the burden for storage management, universal data access with independent geographical locations. The foundations of cloud computing lie in the outsourcing of computing tasks to the third party. It entails the security risks in terms of confidentiality, integrity and availability of data and service. The issue to convince the cloud clients that their data are kept intact is especially vital since the clients do not store these data locally. Remote data integrity checking is a primitive to address this issue. For the general case, when the client stores his data on multicloud servers, the distributed storage and integrity checking are indispensable. On the other hand, the integrity checking protocol must be efficient in order to make it suitable

for capacity-limited end devices. Thus, based on distributed computation, we will study distributed remote data integrity checking model and present the corresponding concrete protocol in multicloud storage.

2. Motivation

In cloud computing, remote data integrity checking is an important security problem. The clients' massive data is outside his control. The malicious cloud server may corrupt the clients' data in order to gain more benefits. In PKI (public key infrastructure), provable data possession protocol needs public key certificate distribution and management. It will incur considerable overheads since the verifier will check the certificate when it checks the remote data integrity. In addition to the heavy certificate verification, the system also suffers from the other complicated certificates management such as certificates generation, delivery, revocation, renewals, etc. In cloud computing, most verifiers only have low computation capacity. Identity-based public key cryptography can eliminate the complicated certificate management.

3. Related Work

In cloud computing, remote data integrity checking is an important security problem. The clients' massive data is outside his control. The malicious cloud server may corrupt the clients' data in order to gain more benefits. Many researchers proposed the corresponding system model and security model. In 2007, provable data possession (PDP) paradigm was proposed by Ateniese et al. [1]. In the PDP data integrity with a high probability. Based on the RSA, they designed two provably secure PDP schemes. After that, Ateniese et al. proposed dynamic PDP model and concrete scheme [2] although it does not support insert operation. In order to support the insert operation, in 2009, Erway et al. proposed a full-dynamic PDP scheme based on the authenticated flip table [3]. The similar work has also been done by F. Sebe et al. [4]. PDP allows a verifier to verify the remote data integrity without retrieving or downloading the whole data. It is a

probabilistic proof of possession by sampling random set of blocks from the server, which drastically reduces I/O costs. The verifier only maintains small metadata to perform the integrity checking. PDP is an interesting remote data integrity checking model. In 2012, Wang proposed the security model and concrete scheme of proxy PDP in public clouds [5]. At the same time, Zhu et al. proposed the cooperative PDP in the multicloud storage [6].

Following Ateniese et al.'s pioneering work, many remote data integrity checking models and protocols have been proposed [7], [8], [9], [10], [11], [12]. In 2008, Shacham presented the first proof of retrievability (POR) scheme with provable security [13]. In POR, the verifier can check the remote data integrity and retrieve the remote data at any time. The state of the art can be found in [14], [15], [16], [17]. On some cases, the client may delegate the remote data integrity checking task to the third party. It results in the third party auditing in cloud computing [18], [19], [20], [21]. One of benefits of cloud storage is to enable universal data access with independent geographical locations. This implies that the end devices may be mobile and limited in computation and storage. Efficient integrity checking protocols are more suitable for cloud clients equipped with mobile end devices.

4. Proposed System

In MultiCloud environment, remote data integrity checking is required to secure user's model, the verifier can check remote data. User will upload file to Cloud. This file is split into blocks using Dynamic Block generation Algorithm and stored in a MultiCloud environment. File Allocation Table (FAT) File System has proper Indexing and Metadata's for the different Chunks of the Cloud Storage. If Attacker corrupts data in MultiCloud, the continuous auditing process helps the Verifier to perform Block level and File level checking for remote data Integrity Checking using Verifiable Data Integrity Checking Algorithm. Cloud provides random blocks to Verifier for Integrity Checking which is to protect user privacy from Verifier (Third Party). File recovery is done by the Verifier automatically if the data gets corrupted during checking. User can complaint Cloud for File Recovery. At every access of the file by the user, blocks of the data will be dynamically reallocated between the Cloud Servers. This achieves access confidentiality in Cloud Computing.

5. Proposed Methodology

5.1 Admin Configuration

Admin configure MultiCloud server setup. Server IP Address and Port number is given by the admin for each Cloud. Now a Server Architecture is created for MultiCloud Storage. If the admin has to reconfigure the old MultiCloud server setup, it can be done. For old server setup, FAT file can be modified or remain same. Audit time will be set by the admin for Data Integrity checking process[22-26].

5.2 Data Upload and Block Split

User has an initial level Registration Process at the web end. The users provide their own personal information for this process. The server in turn stores the information in its database. After Registration, user can upload files to the server. Uploaded files will be stored in a Server. When the user upload the data to different cloud by the time it is Split into different blocks using Dynamic block generation Algorithm and each block will be appended with Signatures before Storing the data in FATFS.

Signature generated using MD5 Algorithm. Also the data gets encoded using for Base64 Algorithm[27].

5.3 Verifier Integrity Checking

FAT FS has proper Indexing and MetaData's for the different Chunks of the Data that is being uploaded by User. Verifier performs Remote Integrity Checking on Cloud Data. Cloud allocates random combination of all the blocks to the Verifier, instead of the whole file is retrieved during integrity checking. This is to protect user privacy from a third party (Verifier). Verifiable Data Integrity Checking Algorithm is done in two steps: Block Checking and File Checking. In Block Checking step: Three signatures are generated for Block level Checking[33-35].

1. A signature of a block retrieved from a FATFS
2. A new signature is generated for block to be checked
3. A Signature is retrieved from the block appended with the signature which is stored in the Cloud The above three signatures are cross checked for Block level Integrity Checking. And the block contents are appended to verify with File level Integrity checking.

5.4 Automatic and On Demand File Recovery

Attacker can corrupt data in any one of the cloud servers. On Data Integrity Checking done by the Verifier, Verifier informs Corrupted blocks to the Cloud. Recovery Process will be done by the verifier automatically when data gets corrupted. User can complaint to the Cloud if the user file get corrupted (Verifier doesn't perform checking on this file). Whenever user access file, Blocks will be reallocated

dynamically to provide access confidentiality in cloud and FAT File System will get updated[36].

6. Conclusion

In multicloud storage, this paper formalizes the ID-DPDP system model and security model. At the same time, we propose the first ID-DPDP protocol which is provably secure under the assumption that the CDH problem is hard. Besides of the elimination of certificate management, our ID-DPDP protocol has also flexibility and high efficiency. At the same time, the proposed ID-DPDP protocol can realize private verification, delegated verification and public verification based on the client's authorization.

References

- [1] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," in Proc. CCS, 2007, pp. 598-609.
- [2] G. Ateniese, R. DiPietro, L.V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," in Proc. SecureComm, 2008, pp. 1-10.
- [3] C.C. Erway, A. Kupcu, C. Papamanthou, and R. Tamassia, "Dynamic Provable Data Possession," in Proc. CCS, 2009, pp. 213-222.
- [4] F. Sebe, J. Domingo-Ferrer, A. Martınez-Balleste, Y. Deswarte, and J. Quisquater, "Efficient Remote Data Integrity Checking in Critical Information Infrastructures," IEEE Trans. Knowl. Data Eng., vol. 20, no. 8, pp. 1034-1038, Aug. 2008.
- [5] H.Q. Wang. (2013, Oct./Dec.). Proxy Provable Data Possession in Public Clouds. IEEE Trans. Serv. Comput. [Online]. (4), pp. 551- 559. Available: <http://doi.ieeecomputersociety.org/10.1109/TSC.2012.35>.
- [6] Y. Zhu, H. Hu, G.J. Ahn, and M. Yu, "Cooperative Provable Data Possession for Integrity Verification in Multicloud Storage," IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 12, pp. 2231-2244, Dec. 2012.
- [7] Y. Zhu, H. Wang, Z. Hu, G.J. Ahn, H. Hu, and S.S. Yau, "Efficient Provable Data Possession for Hybrid Clouds," in Proc. CCS, 2010, pp. 756-758.
- [8] R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "MR-PDP: Multiple-Replica Provable Data Possession," in Proc. ICDCS, 2008, pp. 411-420.
- [9] A.F. Barsoum and M.A. Hasan, "Provable possession and replication of data over cloud servers," Centre Appl. Cryptogr. Res., Univ. Waterloo, Waterloo, ON, Canada, Rep. 2010/32. [Online]. Available: <http://www.cacr.math.uwaterloo.ca/techreports/2010/cacr2010-32.pdf>.
- [10] Z. Hao and N. Yu, "A Multiple-Replica Remote Data Possession Checking Protocol with Public Verifiability," in Proc. 2nd Int. Symp. Data, Privacy, E-Comm., 2010, pp. 84-89.
- [11] A.F. Barsoum and M.A. Hasan, "On verifying dynamic multiple data copies over cloud servers," Int. Assoc. Cryptol. Res., New York, NY, USA, IACR eprint Rep. 447, 2011. [Online]. Available: <http://eprint.iacr.org/2011/447.pdf>.
- [12] A. Juels and B.S. Kaliski, Jr., "PORs: Proofs of Retrievability for Large Files," in Proc. CCS, 2007, pp. 584-597.
- [13] H. Shacham and B. Waters, "Compact Proofs of Retrievability," in Proc. ASIACRYPT, vol. 5350, LNCS, 2008, pp. 90-107.
- [14] K.D. Bowers, A. Juels, and A. Oprea, "Proofs of Retrievability: Theory and Implementation," in Proc. CCSW, 2009, pp. 43-54.
- [15] Q. Zheng and S. Xu, "Fair and Dynamic Proofs of Retrievability," in Proc. CODASPY, 2011, pp. 237-248.
- [16] Y. Dodis, S. Vadhan, and D. Wichs, "Proofs of Retrievability via Hardness Amplification," in Proc. TCC, vol. 5444, LNCS, 2009, pp. 109-127.
- [17] Y. Zhu, H. Wang, Z. Hu, G.J. Ahn, and H. Hu, "Zero-Knowledge Proofs of Retrievability," Sci. Chin. Inf. Sci., vol. 54, no.8, pp. 1608-1617, Aug. 2011.
- [18] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing," in Proc. IEEE INFOCOM, Mar. 2010.
- [19] 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 Distrib. Syst., vol. 22, no. 5, pp. 847-859, May 2011.
- [20] C. Wang, Q. Wang, K. Ren, N. Cao, and W. Lou, "Toward Secure and Dependable Storage Services in Cloud Computing," IEEE Trans. Serv. Comput., vol. 5, no. 2, pp. 220-232, Apr./June 2012.
- [21] Udayakumar R., Kaliyamurthi K.P., Khanaa, Thooyamani K.P., Data mining a boon: Predictive system for university topper women in academia, World Applied Sciences Journal, v-29, i-14, pp-86-90, 2014.
- [22] Kaliyamurthi K.P., Parameswari D., Udayakumar R., QOS aware privacy preserving location monitoring in wireless sensor network, Indian Journal of Science and Technology, v-6, i-SUPPL5, pp-4648-4652, 2013.
- [23] Brintha Rajakumari S., Nalini C., An efficient cost model for data storage with horizontal layout in the cloud, Indian Journal of Science and Technology, v-7, i-, pp-45-46, 2014.

- [24] Brintha Rajakumari S., Nalini C., An efficient data mining dataset preparation using aggregation in relational database, *Indian Journal of Science and Technology*, v-7, i-, pp-44-46, 2014.
- [25] Khanna V., Mohanta K., Saravanan T., Recovery of link quality degradation in wireless mesh networks, *Indian Journal of Science and Technology*, v-6, i-SUPPL.6, pp-4837-4843, 2013.
- [26] Khanaa V., Thooyamani K.P., Udayakumar R., A secure and efficient authentication system for distributed wireless sensor network, *World Applied Sciences Journal*, v-29, i-14, pp-304-308, 2014.
- [27] Udayakumar R., Khanaa V., Saravanan T., Saritha G., Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction, *Middle - East Journal of Scientific Research*, v-16, i-12, pp-1781-1785, 2013.
- [28] Khanaa V., Mohanta K., Saravanan. T., Performance analysis of FTTH using GEAPON in direct and external modulation, *Indian Journal of Science and Technology*, v-6, i-SUPPL.6, pp-4848-4852, 2013.
- [29] Kaliyamurthi K.P., Udayakumar R., Parameswari D., Mugunthan S.N., Highly secured online voting system over network, *Indian Journal of Science and Technology*, v-6, i-SUPPL.6, pp-4831-4836, 2013.
- [30] Thooyamani K.P., Khanaa V., Udayakumar R., Efficiently measuring denial of service attacks using appropriate metrics, *Middle - East Journal of Scientific Research*, v-20, i-12, pp-2464-2470, 2014.
- [31] R.Kalaiprasath, R.Elankavi, Dr.R.Udayakumar, Cloud Information Accountability (Cia) Framework Ensuring Accountability Of Data In Cloud And Security In End To End Process In Cloud Terminology, *International Journal Of Civil Engineering And Technology (Ijciet)* Volume 8, Issue 4, Pp. 376–385, April 2017.
- [32] R.Elankavi, R.Kalaiprasath, Dr.R.Udayakumar, A fast clustering algorithm for high-dimensional data, *International Journal Of Civil Engineering And Technology (Ijciet)*, Volume 8, Issue 5, Pp. 1220–1227, May 2017.
- [33] R. Kalaiprasath, R. Elankavi and Dr. R. Udayakumar. Cloud. Security and Compliance - A Semantic Approach in End to End Security, *International Journal Of Mechanical Engineering And Technology (Ijmet)*, Volume 8, Issue 5, pp-987-994, May 2017.
- [34] Thooyamani K.P., Khanaa V., Udayakumar R., Virtual instrumentation based process of agriculture by automation, *Middle - East Journal of Scientific Research*, v-20, i-12, pp-2604-2612, 2014.
- [35] Udayakumar R., Thooyamani K.P., Khanaa, Random projection based data perturbation using geometric transformation, *World Applied Sciences Journal*, v-29, i-14, pp-19-24, 2014.
- [36] Udayakumar R., Thooyamani K.P., Khanaa, Deploying site-to-site VPN connectivity: MPLS Vs IPSec, *World Applied Sciences Journal*, v-29, i-14, pp-6-10, 2014.
- [37] DR.P.SHUNMUGAPRIYA, R.MARAGATHAM, “Fidoop-Dp: Data Partitioning In Frequent Itemset Mining On Hadoop Clust”, *International Innovative Research Journal of Engineering and Technology*, 2017.

