



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 6, June 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.165



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Heterogeneous Data Storage Management with Deduplication in Cloud Computing

Nithish Kumar, Dr. V.P. Gladis Pushaparathi

U.G Scholar, Department of CSE, Velammal Institute of Technology, Panchetti, Tamilnadu, India

Professor, Department of CSE, Velammal Institute of Technology, Panchetti, Tamilnadu, India

ABSTRACT: Cloud storage as one of the most important services of cloud computing. Data ownership proof is an essential process of data deduplication, especially for encrypted data. But this scheme does not provide flexible deduplication control across multiple Cloud Service Providers (CSPs). In this paper, They propose a multiple cloud service provider (CSPs) in which the data owner will upload the file and the hash MD5 algorithm is used to check data duplication during data storage at the cloud. CSPs. It can achieve data deduplication and access control with different security requirements. And also they have proposed a scheme called Provable Ownership of the File (POF). The result it is security, effectiveness and efficiency towards data storage management.

I. INTRODUCTION

Cloud computing allows centralized data storage and online access to computer services or resources. It offers a new way of Information Technology (IT) services by rearranging various resources and providing them to users based on their demands. Cloud computing has greatly enriched pervasive services and become a promising service platform due to a number of desirable properties, such as scalability, elasticity, fault-tolerance, and pay-per-use. Data storage service is one of the most widely consumed cloud services. Cloud users have greatly benefited from cloud storage since they can store huge volume of data without upgrading their devices and access them at any time and in any place. However, cloud data storage offered by Cloud Service Providers (CSPs) still incurs some problems.

First of all, various data stored at the cloud may request different ways of protection due to different data sensitivity. The data stored at the cloud include sensitive personal information, publicly shared data, data shared within a group, and so on. Obviously, crucial data should be protected at the cloud to prevent from any access of unauthorized parties. Some important data, however, have no such a requirement. As outsourced data could disclose personal or even sensitive information, data owners sometimes would like to control their data by themselves, while on some occasion, they prefer to delegate their control to a third party since they cannot be always online or have no idea how to perform such a control. How to make cloud data access control adapt to various scenarios and satisfy different user demands becomes a practically important issue. Access control on encrypted data has been widely studied in the literature.

II. LITERATURE SURVEY

Zheng Yan (M'06, SM'14) received the BEng degree in electrical engineering and the MEng degree in computer science and engineering from the Xi'an Jiaotong University, Xi'an, China, in 1994 and 1997, respectively, the second MEng degree in information security from the National University of Singapore, Singapore, in 2000, and the licentiate of science and the doctor of science in technology in electrical engineering from Helsinki University of Technology, Helsinki, Finland, in 2005 and 2007.

She is currently a professor at the Xidian University, Xi'an, China and a visiting professor with the Aalto University, Espoo, Finland. She authored more than 150 peer-reviewed publications and solely authored two books. She is the inventor and coinventor of about 60 patents and PCT patent applications.

Her research interests are in trust, security and privacy, social networking, cloud computing, networking systems, and data mining. She serves as an associate editor of the Information Sciences, Information Fusion, the IEEE Internet of Things Journal, the IEEE Access Journal, the Journal of Network and Computer Applications, the Security and

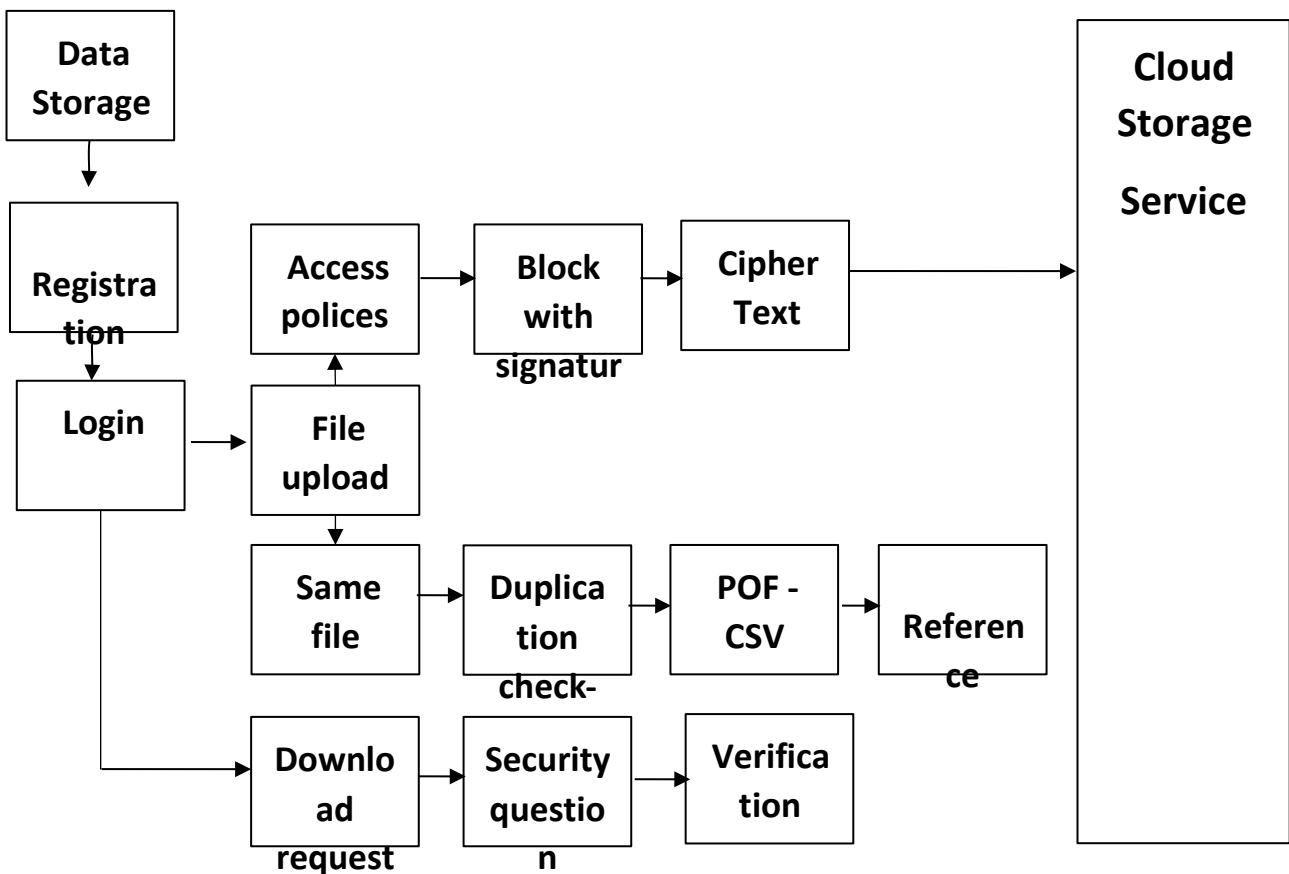


Communication Networks, etc. She is a leading guest editor of many reputable journals including ACM TOMM, FGCS, IEEE Systems Journal, MONET, etc. She served as a steering, organization and program committee member for more than 70 international conferences. She is a senior member of the IEEE

Wenxiu Ding received the BEng degree in information security from the Xidian University, Xi'an, China, in 2012. She is currently working toward the PhD degree in information security, the School of Cyber Engineering in the Xidian University, and also a visiting research student at the School of Information

III. PROPOSED SYSTEM

In this work, we proposed to storage across multiple CSP's and preserve data security by managing deduplication. we also introduced a scheme called Provable Ownership of the File(POF).They enhance user privacy and improve the performance of practical deployment. The random hash code challenge is applied to verify data ownership, which can guarantee that the data holder really have the original data rather than its hash code



CLOUD USER AUTHENTICATION

Owner has an initial level Registration Process in Cloud Service Provider(CSP). The users provide their own personal information for this process. The server in turn stores the information in its database. Then they have the Login process for the further access in cloud service Provider.

FILE UPLOAD AND COMPARISON

In this module, the data Owner create their account under the Public cloud and upload the file in cloud storage. Here the Provable Ownership of the file(POF) scheme is proposed. While uploading the file by data owner, the hash key is generated based on MD5 algorithm. The hash key is unique for all the upload files. But if the same file is upload by the other data Owner it will not allow the file to upload rather then it will replace the reference id through Mapping of index. It also check the file for physical present or not by both the data Owner.

SET ACCESS POLICY FOR FILE

In this module User will chooses the file and uploads to Storage where the HDFS storage system .In the system will generate a signature in particular file and then split into multiple block. Each block will be generate signature with key . In the signature by using MD5 message-digest algorithm is cryptographic hash function producing a 128-bit hash value typically expressed in text format as 32 digit hex value so that files of same are de-duplicated. After that generate convergent keys for each blocks splitting to store CSV file .like filename, file path, blocks, username, password and block keys.

FILE DOWNLOAD REQUEST AND HANDLING

In this module , the data owner will download the file from cloud service provider. If they do not find the file then they will request to download the file from different Cloud service provider and also check whether the file is present or not then it gives the response to data owner.

IV. RESULT

Data deduplication is important and significant in the practice of cloud data storage, especially for big data storage management. In this paper, we proposed a heterogeneous data storage management scheme, which offers flexible cloud data deduplication and access control. Our scheme can adapt to various application scenarios and demands and offer economic big data storage management across multiple CSPs. It can achieve data deduplication and access control with different security requirements. Security analysis, comparison with existing work and implementation based performance evaluation showed that our scheme is secure, advanced and efficient.

V. CONCLUSION

Our scheme supports data privacy of cloud users since the data stored at the cloud is in an encrypted form. One way to support identity privacy is to apply pseudonyms in Key Generation Centre (KGC), where a real identity is linked to a number of pseudonyms, which is verified and certified by the KGC. In our future work, we will further enhance user privacy and improve the performance of our scheme towards practical deployment. In addition, we will conduct game theoretical analysis to further prove the rationality and security of the proposed scheme.

REFERENCES

- [1] R. Chow, et al., "Controlling data in the cloud: outsourcing computation without outsourcing control," in Proc. ACM Workshop Cloud Computer. Secure., 2009, pp. 85–90.
- [2] S. Kamara, and K. Lauter, "Cryptographic cloud storage," Finance. Crypto. Data Secure., 2010, pp. 136–149.
- [3] Q. Liu, C. C. Tan, J. Wu, and G. Wang, "Efficient information retrieval for ranked queries in cost-effective cloud environments," in Proc. IEEE INFOCOM, 2012, pp. 2581–2585.
- [4] M. Kallahalla, E. Riedel, R. Swaminathan, Q. Wang, and K. Fu, Plutus: scalable secure file sharing on untrusted storage," in Proc. USENIX Conf. File Storage Technol., 2003, pp. 29–42.
- [5] E.-J. Goh, H. Shacham, N. Modadugu, and D. Boneh, "SiRiUS: securing remote untrusted storage," in Proc. Netw. Distrib. Syst. Secur. Symp., 2003, pp. 131–145.
- [6] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in Proc. IEEE Symp. Secur. Privacy, 2007, pp. 321–334.
- [7] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in Proc. 13th ACM Comput. Communication. Secure., 2006, pp. 89–98.



- [8] S. Muller, S. Katzenbeisser, and C. Eckert, "Distributed attribute based encryption," in Proc. 11th Annu. Int. Conf. Inf. Secur. Crypto., 2008, pp. 20–36.
- [9] A. Sahai, and B. Waters, "Fuzzy identity-based encryption," in Proc. 24th Int. Conf. Theory App. Cryptographic Tech., 2005, pp. 457–473.
- [10] S. C. Yu, C. Wang, K. Ren, and W. J. Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing," in Proc. IEEE INFOCOM, 2010, pp. 534–542.
- [11] G. J. Wang, Q. Liu, J. Wu, and M. Y. Guo, "Hierarchical attribute based encryption and scalable user revocation for sharing data in cloud servers," Computer. Secur., vol. 30, no. 5, pp. 320–331, 2011.
- [12] S. C. Yu, C. Wang, K. Ren, and W. J. Lou, "Attribute based data sharing with attribute revocation," in Proc. ACM Asia Conf. Comput. Commun. Secur., 2010, pp. 261–270.
- [13] G. J. Wang, Q. Liu, and J. Wu, "Hierarchical attribute-based encryption for fine-grained access control in cloud storage services," in Proc. 17th ACM Comput. Commun. Secur., 2010, pp. 735–737.
- [14] M. Zhou, Y. Mu, W. Susilo, M. H. Au, and J. Yan, "Privacy preserved access control for cloud computing," in Proc. IEEE 10th
- [15] Z. G. Wan, J. E. Liu, and R. H. Deng, "HASBE: A hierarchical attributebased solution for flexible and scalable access control in cloud computing," IEEE Trans. Inf. Forensics Secur., vol. 7, no. 2, pp. 743–754, Art. 2012.
- [16] Z. Yan, Trust Management in Mobile Environments – Usable and Autonomic Models. Hershey, PA, USA: IGI Global, 2013.
- [17] Y. Tang, P. P. Lee, J. C. Lui, and R. Perlman, "Secure overlay cloud storage with access control and assured deletion," IEEE Trans. Dependable Secure Comput., vol. 9, no. 6, pp. 903–916, Nov.-Dec. 2012. [18] M. Bellare, S. Keelveedhi, and T. Ristenpart, "DupLESS: server aided encryption for deduplicated storage," in Proc. 22nd USENIX Conf. Secur., 2013, pp. 179–194. [19] Dropbox, "A file-storage and sharing service." [Online]. Available: <http://www.dropbox.com/>, retrieved March 2017.
- [20] Google Drive. [Online]. Available: <http://drive.google.com>, retrieved May 2017.
- [21] Mozy, "Mozy: A file-storage and sharing Service." [Online]. Available: <http://mozy.com/>, retrieved May 2017.
- [22] J. R. Douceur, A. Adya, W. J. Bolosky, P. Simon, and M. Theimer, "Reclaiming space from duplicate files in a serverless distributed file system," in Proc. 22nd Int. Conf. Distributed Comput. Syst., 2002, pp. 617–624. pp. 1615–1625, Jun. 2014.

BIOGRAPHY

NITHISH KUMAR K is a B.E final year student in the department of computer science engineering from velammal institute of technology. His current research is focuses on HETEROGENEOUSDATA STORAGE MANAGEMENT WITH DEPLICATION IN CLOUD COMPUTING

Dr. V.P. GLADIS PUSHPARATHI professor in the department of computer science engineering and project coordinator in the department of computer science engineering velammal institute of technology.



INNO  SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

 **doi**[®]
cross **ref**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details