IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-6, Special Issue-2, March 2018 E-ISSN: 2347-2693

Avoid Deduplication on Cloud through ABE

[1]P.Parameshwari, [2]S.Padmapriya,[3]G.Lakshmipriya

[1]Department of Computer Science, Srimad Andavan Arts and Science College (Autonomous), Trichy-620005
[2]Department of Computer Science, Srimad Andavan Arts and Science College (Autonomous), Trichy-620005
[3] Department of Computer Science, Srimad Andavan Arts and Science College (Autonomous), Trichy-620005

Available online at: www.ijcseonline.org

Abstract- Attribute-based encryption has been widely second-hand in cloud computing where a data provider outsources his/her encrypted data to a blur service provider, and can split the data with users possessing specific qualifications. However, the standard ABE scheme does not support protected deduplication, which is crucial for eliminating duplicate copies of the same data in arrange to save cargo space space and system bandwidth. Here, present an attribute-based storage system with protected deduplication in a hybrid cloud setting, where a confidential cloud is in charge for duplicate uncovering and a public blur manages the storage. Compared with the prior data deduplication systems, our scheme has two advantages. Firstly, it can be second-hand to in secret share data by means of users by specifying admission policies rather than distribution decryption keys. Secondly, it achieves the standard view of semantic refuge for data discretion while existing systems only attain it by defining a weaker refuge notion. In addition, put forth a method to modify a ciphertext over one admission policy into ciphertexts of the same plaintext but beneath other right of entry policies without enlightening the underlying plaintext.

Keywords- Cloud Computing, ABE, Architecture

I. INTRODUCTION

Cloud computing greatly facilitates data providers who desire to outsource their data to the cloud without disclosing their sensitive data to outside parties and would like users with certain qualifications to be able to access the data . This requires data to be stored in encrypted forms with admission control policies such that no one apart from users with attributes (or credentials) of precise forms can decrypt the encrypted data. An encryption method that meets this requirement is called attribute-based encryption, where a user's private key is linked with an attribute set, a communication is encrypted beneath an access policy (or access structure) over a set of attributes, and a user can decrypt a ciphertext with his/her personal key if his/her set of attributes satisfies the admission policy associated through this ciphertext. However, the standard ABE scheme fails to achieve protected deduplication.On the other hand, to the best of our information, existing constructions for secure deduplication are not built on attribute-based encryption.

Nevertheless, since ABE and secure deduplication have been extensively practical in cloud computing, it would be attractive to design a cloud storage arrangement possessing both properties. consider the following situation in the design of an attribute-based storage system supporting protected deduplication of encrypted data in the cloud, in which the cloud will not store a folder more than once even although it may receive manifold copies of the similar file encrypted under dissimilar access policies.

In a characteristic storage system with secure deduplication, to accumulate a file in the cloud, a data supplier generates a tag and a ciphertext. The data supplier uploads the tag and the ciphertext to the cloud. Upon in receipt of an outsourcing request from a data supplier for uploading a ciphertext and an linked tag, the cloud runs a so-called parity checking algorithm, which checks if the tag in the inward request is the same to any tags in the storage system. If there is a match, then the causal plaintext of this received ciphertext has already been stored and the new ciphertext is discarded.

It is obvious that such a scheme with a tag appended to the ciphertext does not give the normal notion of semantic safety for information privacy, because if the plaintexts can be predicated from their tags, an adversary can forever make a correct guess by computing the tag of a plaintext and then difficult it against the tag in the confront phase in the semantic safety game. To circumvent this obstruction, bring in our system a cross cloud architecture, which consists of a private cloud accountable for tag checking and ciphertext regeneration and a free cloud storing the ciphertexts.

However, endowing such a tag examination ability to the confidential cloud is not enough to achieve deduplication in the attribute-based storage scheme which employs CP-ABE for data encryption. In the planned attributed-based system, the similar file could be encrypted to dissimilar ciphertexts associated with different access policies, storing only one ciphertext of the file means that users whose attributes satisfy the right to use policy of a unnecessary ciphertext will be deprived of to access the data that they are at liberty to. Another key defy in safe deduplication is to make it protected against duplicate faking attacks in which a legally generated message is imperceptibly replaced by a fake one.

II. RELATED WORK

[1] Cloud storage services are increasingly used by consumers, business, and government. These services are fairly easy to obtain. Google Drive is a popular service, providing users a costeffective, and in some cases free, ability to access, store, collaborate, and disseminate data. It is difficult to identify, acquire, and preserve the evidences when criminals use disparate services. This study was undertaken to determine the data remnants on a Windows computer. We focus on exploring the cloud activities of Google Drive and try to obtain evidences that may be left under these activities, different Internet browsers. By determining the data remnants on client devices, we attempt to enhance the efficiency of the digital forensics and crime investigation

[2] Disk based storage has emerged as the new production storage space categorization for scheme data resistance to replace tape libraries. Deduplication removes superfluous data segments to squeeze data into a extremely compact form and makes it inexpensive to store backups on disk instead of tape. A crucial obligation for venture data defense is high throughput, characteristically over 100 MB/sec, which enables backups to total quickly. A important challenge is to recognize and eradicate copy data segments at this rate on a low-cost scheme that cannot pay for enough RAM to store an index of the stored segments and may be compulsory to right of entry an on-disk index for every effort segment. This paper describes three techniques employed in the manufacture Data Domain deduplication file scheme to ease the disk bottleneck. These techniques include: (1) the Summary Vector, a dense in-memory data structure for identifying new segments; (2) Locality sealed Caching, which maintains the area of the fingerprints of duplicate segments to attain high store hit ratios. Together, they can remove 99% of the disk accesses for deduplication of real world workloads. These techniques allow a modern two-socket dual-core organization to run at 90% CPU use with only one shelf of 15 disks and attain 100 MB/sec for single-stream throughput and 210 MB/sec for multi-stream throughput.

[3] a novel cryptographic early on message locked encryption, where the payment under which encryption and decryption are performed is itself ensuing from the memorandum. MLE provides a way to attain protected deduplication, a goal at there targeted by common cloud-storage providers. To provide definitions both for time unaccompanied and for a look of honesty that call tag constancy. Based on this foundation, make both helpful and hypothetical contributions. On the sensible side, provide ROM security analyses of a usual family of MLE schemes that includes deployed schemes. On the academic side the begin is standard replica solutions, and make relations with deterministic encryption, hash functions secure on linked inputs and the example then take out example to bring scheme under dissimilar assumptions and for dissimilar lessons of memorandum sources.

[4] . Motivated by the difficulty of avoiding duplication in storage systems have lately put forward the idea of Message-Locked Encryption schemes which subsumes convergent encryption and its variants. Such schemes do not rely on enduring secret keys, but quite encrypt messages by keys derived from the communication themselves. To make stronger the notions of security by bearing in mind plaintext distributions that may depend on the community parameters of the schemes. It refers to such inputs as lock-dependent messages. To build two schemes that satisfy new ideas of security for message-locked encryption with lock-dependent messages. main construction deviates from the move towards by avoiding the use of ciphertext components resulting deterministically from the messages. The design a fully randomized scheme that supports an equality-testing algorithm distinct on the ciphertexts. Our subsequent edifice has a deterministic ciphertext part that enables more well-organized parity testing. Security for lock-dependent mail still holds beneath computational assumptions on the communication distributions shaped by the attacker. In both of our schemes the overhead in the length of the ciphertext is only preservative and independent of the communication length.

[5] Cloud storage service providers, and others execute deduplication to keep space by only storing one reproduction of each file uploaded. Should clients conventionally encrypt their files, investments are lost. Message locked encryption . However it is essentially topic to brute force attacks that can get healthy files decrease into a known set. To propose construction that provides protected deduplicated storage resisting creature force attacks, and comprehend it in a organization called DupLESS. In DupLESS, clients encrypt under message based keys obtained from a key-server via an unaware PRF protocol. It enables customers to store encrypted data with an available service, have the renovate perform deduplication on their behalf, and yet achieves strong privacy guarantees. To show that encryption for deduplicated storage space can attain appearance and space savings close to that of using the storage space fix with plaintext data.

III. METHODOLOGY

EXISING PROCESS

Cloud computing greatly facilitates data providers who want to outsource their data to the cloud without disclosing their sensitive data to external parties and would like users with certain credentials to be able to access the data. This requires data to be stored in encrypted forms with access control policies such that no one except users with attributes (or credentials) of specific forms can decrypt the encrypted data. An encryption technique that meets this requirement is called attribute-based encryption, where a user's private key is associated with an attribute set, a message is encrypted under an access policy (or access structure) over a set of attributes, and a user can decrypt a cipher text with his/her private key if his/her set of attributes satisfies the access policy associated with this cipher text.

PROPOSED PROCESS

Firstly, the system is the first that achieves the normal notion of semantic safety for data privacy in attribute-based deduplication systems by resorting to the cross cloud structural design. Secondly, put forth a method to modify a ciphertext over one right of entry policy into cipher texts of the similar plaintext but under any additional access policies without revealing the fundamental plaintext. This technique strength be of self-governing interest in adding to the application in the projected storage system. Thirdly, propose an move toward based on two cryptographic primitives, including a zero-knowledge proof of acquaintance and a promise scheme , to achieve data steadiness in the system.

ARCHITECTURE

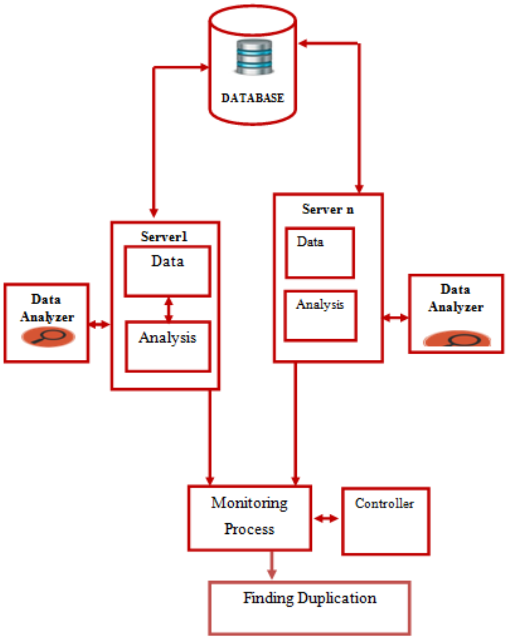

Fig 1: Architecture

IV. CONCLUSION

Attribute-based encryption has been extensively used in cloud computing where information providers subcontract their encrypted data to the cloud and can split the data with users possessing particular credentials. On the other hand, deduplication is an significant technique to save the storage space space and system bandwidth, which eliminates duplicate copies of the same data. However, the standard ABE systems do not hold up secure deduplication, which makes them expensive to be applied in some profitable storage services. In this obtainable a novel approach to comprehend an attribute-based storage space system behind secure deduplication. Our storage system is built under a hybrid cloud architecture, where a private cloud manipulates the calculation and a public cloud manages the storage. The confidential cloud is provided with a trapdoor key associated with the matching ciphertext, with which it can transfer the ciphertext over one admission policy into ciphertexts of the same plaintext under any other right of entry policies without being conscious of the underlying plaintext. It achieves the standard notion of semantic security while existing deduplication schemes only achieve it under a weaker security notion.

PERFORMANCE ANALYSIS

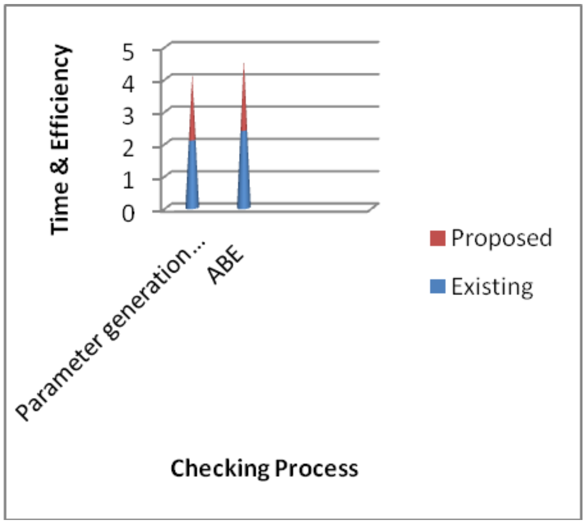

Fig 2: Performance Analysis

REFERENCES

[1] D. Quick, B. Martini, and K. R. Choo, Cloud Storage Forensics. Syngress Publishing / Elsevier, 2014. [Online]. Available: http://www.elsevier.com/books/cloud-storageforensics/ quick/978-0-12-419970-5

[2] K. R. Choo, J. Domingo-Ferrer, and L. Zhang, “Cloud cryptography: Theory, practice and future research directions,” Future Generation Comp. Syst., vol. 62, pp. 51–53, 2016.

[3] K. R. Choo, M. Herman, M. Iorga, and B. Martini, “Cloud forensics: State-of-the-art and future directions,” Digital Investigation, vol. 18, pp. 77–78, 2016.

[4] Y. Yang, H. Zhu, H. Lu, J.Weng, Y. Zhang, and K. R. Choo, “Cloud based data sharing with fine-grained proxy re-encryption,” Pervasive and Mobile Computing, vol. 28, pp. 122–134, 2016.

[5] D. Quick and K. R. Choo, “Google drive: Forensic analysis of data remnants,” J. Network and Computer Applications, vol. 40, pp. 179– 193, 2014.

[6] A. Sahai and B. Waters, “Fuzzy identity-based encryption,” in Advances in Cryptology - EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005, Proceedings, ser. Lecture Notes in Computer Science, vol. 3494. Springer, 2005, pp. 457–473.

[7] B. Zhu, K. Li, and R. H. Patterson, “Avoiding the disk bottleneck in the data domain deduplication file system,” in 6th USENIX Conference on File and Storage Technologies, FAST 2008, February 26- 29, 2008, San Jose, CA, USA. USENIX, 2008, pp. 269–282.

[8] M. Bellare, S. Keelveedhi, and T. Ristenpart, “Message-locked encryption and secure deduplication,” in Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings, ser. Lecture Notes in Computer Science, vol. 7881. Springer, 2013, pp. 296–312.

[9] M. Abadi, D. Boneh, I. Mironov, A. Raghunathan, and G. Segev, “Message-locked encryption for lock-dependent messages,” in Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I, ser. Lecture Notes in Computer Science, vol. 8042. Springer, 2013, pp. 374–391.

[10] S. Keelveedhi, M. Bellare, and T. Ristenpart, “Dupless: Serveraided encryption for deduplicated storage,” in Proceedings of the 22th USENIX Security Symposium, Washington, DC, USA, August 14-16, 2013. USENIX Association, 2013, pp. 179–194.

[11] M. Bellare and S. Keelveedhi, “Interactive message-locked encryption and secure deduplication,” in Public-Key Cryptography - PKC 2015 - 18th IACR International Conference on Practice and Theory in Public-Key Cryptography, Gaithersburg, MD, USA, March 30 - April 1, 2015, Proceedings, ser. Lecture Notes in Computer Science, vol. 9020. Springer, 2015, pp. 516–538.

[12] S. Bugiel, S. N¨ urnberger, A. Sadeghi, and T. Schneider, “Twin clouds: Secure cloud computing with low latency - (full version),” in Communications and Multimedia Security, 12th IFIP TC 6 / TC 11 International Conference, CMS 2011, Ghent, Belgium, October 19- 21,2011. Proceedings, ser. Lecture Notes in Computer Science, vol. 7025. Springer, 2011, pp. 32–44.

[13] S. Goldwasser, S. Micali, and C. Rackoff, “The knowledge complexity of interactive proof-systems (extended abstract),” in Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6-8, 1985, Providence, Rhode Island, USA. ACM, 1985, pp. 291– 304.

[14] M. Fischlin and R. Fischlin, “Efficient non-malleable commitment schemes,” in Advances in Cryptology - CRYPTO 2000, 20th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2000, Proceedings, ser. Lecture Notes in Computer Science, vol. 1880. Springer, 2000, pp. 413–431.

[15] S. Goldwasser and S. Micali, “Probabilistic encryption,” J. Comput. Syst. Sci., vol. 28, no. 2, pp. 270–299, 1984.