International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 07
March 2018

Secure Deduplication of Encrypted Data in Cloud Using Attribute Based Encryption (ABE)

[1]P. Anwesh, [2]M.Divya, [3]D. Sanjay, [4]K. Nayana, [5]Mr. A. Mruthyunjayam

[1,2,3,4]B. Tech Students, Department of Computer Science &Engineering, St. Martin's Engineering College, Dhulapally Village, Qutbullapur Mandal, Secunderabad, Telangana, India.

[5]Associate Professor, Department of Computer Science & Engineering, St. Martin's Engineering College, Dhulapally Village, Qutbullapur Mandal, Secunderabad, Telangana, India.

ABSTRACT—*At present data duplication will be increasing in the cloud storage areas and due to this duplicate content the cloud storage space may reduced. To improve the storage space of the cloud, we need to perform the deduplication on cloud storage space. In this paper, we are implementing Attribute-based Encryption (ABE) scheme used to support the secure deduplication. Not only secure deduplication, in this paper we also implementing access policies to share the data confidentially to the cloud users.In our proposed system the deduplication processes done by the private server not like traditional deduplication schemes. From the experimental results we can prove that the proposed system can significantly improve the secure deduplication performance along with confidentially data sharing.*

Keywords: *Cloud Computing, Data Storage, ABE, Data Deduplication, Access Control*

1. INTRODUCTION

Data, being a previous component on this age of computing, has been serviced with technology in masses from main domains inclusive of garage, transfer and security. Data garage has evolved loads considering the fact that its inception and has scaled up from magnetic tapes to the cloud. The alternate of facts has moreover been advanced to a quicker and cozy way. Security of information has been a high challenge in the area of garage and switch and is being addressed even now. Storing the records has been a chief subject matter while the data available turned into growing in abundance. The advent of cloud garage has addressed the difficulty to a superb volume. Due to the increase in facts usage, the storage area have been compromised for redundant information. Repeated incidence of the same statistics furnished by thousands and thousands of customers creates wastage of available area on the cloud. Deduplication, whilst deployed allows the garage of a report uploaded by means of a client and transparently rejects all in addition uploads with the aid of exclusive customers for the equal file thereby supplying high financial savings in storage space. The document is shared among all the clients who've attempted to upload the equal.

Data deduplication has certain advantages to Eliminating redundant facts can notably cut back storage necessities and increase bandwidth performance. Since number one garage has gotten inexpensive over the years, typically shop many variations of the equal information so that new workers can reuse earlier work finished. Some operations like backup save extremely redundant statistics. Data deduplication is statistics compression approach for casting off reproduction copies of repeating facts in garage. This approach is used to improve storage

utilization and can also be implemented to community information transfers to lower the range of bytes that ought to be sent. Deduplication eliminates redundant information by using retaining simplest one physical replica and referring other redundant records to that copy rather than retaining multiple information copies with the equal content.

Fine-grained access control systems facilitate granting differential get entry to rights to a set of users and permit flexibility in specifying the get right of entry to rights of individual customers. Several techniques are recognized for imposing excellent grained get entry to manipulate. Common to the prevailing strategies and the references therein) is the truth that they hire a depended on server that stores the records in clean. Access control relies on software checks to ensure that a user can get admission to a bit of facts most effective if he's legal to achieve this. This scenario isn't always mainly attractive from a protection viewpoint. In the occasion of server compromise, as an instance, because of a software vulnerability exploit, the capacity for information theft is vast. Furthermore, there is usually a chance of "insider attacks" in which someone gaining access to the server steals and leaks the statistics, for example, for financial profits. Some strategies create consumer hierarchies and require the customers to share a common secret key if they are in a common set in the hierarchy. The statistics is then categorized consistent with the hierarchy and encrypted below the public key of the set it is meant for. Clearly, such methods have numerous barriers. If a 3rd party have to get right of entry to the information for a fixed, a consumer of that set both desires to behave as an intermediary and decrypt all applicable entries for the birthday party or have to deliver the celebration its non-public decryption key, and for this reason permit it have get admission to to all entries. In many instances, by the usage of the user hierarchies it isn't always even viable to understand a get right of entry to manage equal to monotone get admission to timber.

2. LITERATURE SURVEY

Although confident deletion is a vast hurdle for adoption of public clouds, it may also emerge as a differentiator within the marketplace. Allowing cloud users to govern and confirm how their data is dealt with is vital for even greater adoption. K. R. Choo, J. Domingo-Ferrer, and L. Zhang have proven the importance of assuring deletion within the cloud and supplied assured deletion for both the cloud tenant and the company. In instances in which a bent cloud issuer is used, they have surveyed and mentioned present solutions against requirements and outlined their obstacles. For the sincere provider, they have got outlined confident deletion requirements for the company, reviewed contemporary infrastructures then furnished a systematization of assured deletion demanding situations their features pose almost about confident deletion.

Y. Yang, H. Zhu, H. Lu, J. Weng, Y. Zhang, and K. R. Choo proposed a new single-hop unidirectional CP-ABPRE scheme, which supports characteristic-based totally re-encryption with any monotonic get right of entry to shape, to address the open hassle left by way of the existing CP-ABPRE schemes. They additionally confirmed that our scheme may be proved IND-sAS-CCA secure inside the random oracle version assuming the decisional q-parallel BDHE assumption holds.

R. Ostrovsky, A. Sahai, and B. Waters provided the primary Attribute-Based Encryption gadget that supports the expression of non-monotone formulas in key guidelines. They carried out this thru a unique utility of revocation methods into present ABE schemes. In

addition, the performance of our scheme compares very favorably to that of present, much less-expressive ABE structures. An essential goal in ABE structures is to create even more expressive systems. They paintings took a massive step forward by way of allowing key policies which could explicit any get right of entry to system. Eventually, they would like to have systems that could explicit any access circuit.

V. Goyal, A. Jain, O. Pandey, and A. Sahai first of all studied the possible revocation operations in CP-ABE scheme: unmarried characteristic revocation, attribute set revocation and particular identifier revocation. Then, based on specific identifier revocation method, they proposed the CP-ABE-R scheme in which malicious users may be efficiently revoked. They presented the ciphertext coverage characteristic primarily based encryption scheme with efficient revocation by way of using linear secret sharing scheme and binary tree approach as the underlying equipment. They have proven that the delegating capability may be without problems provided in the proposed scheme, but all the delegatees are related to their unique delegator's particular identifier.

3. OVERVIEW OF THE SYSTEM

A. System Overview

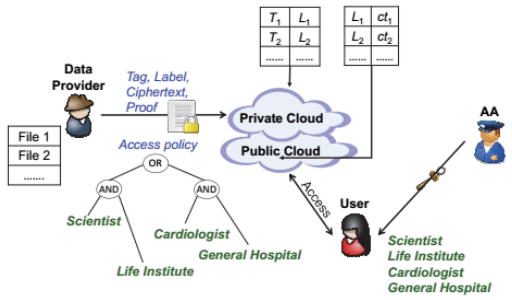


Fig1. System Framework

We offered a singular method to realize an characteristic-primarily based garage system helping comfy deduplication. Our garage device is built under a hybrid cloud architecture, in which a non-public cloud manipulates the computation and a public cloud manages the storage. The personal cloud is provided with a trapdoor key related to the corresponding ciphertext, with which it may transfer the ciphertext over one get right of entry to policy into ciphertexts of the identical plaintext beneath every other access guidelines without being aware about the underlying plaintext. After receiving a storage request, the non-public cloud first exams the validity of the uploaded item thru the attached proof.

B. Share Data Confidentiality by Access Control

Unauthorized users must be avoided from getting access to the plaintext of the shared facts saved in the cloud server. In addition, the cloud server, which is supposed to be honest however curious, ought to also be deterred from understanding plaintext of the shared data. In this proposed system, the data owner can provide the access control on the registered users.

4. METHODOLGY

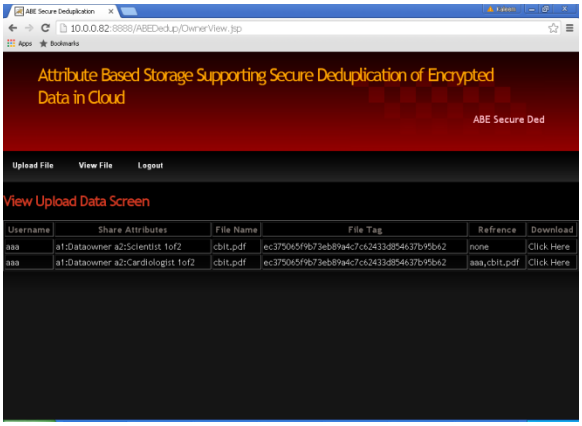
Secure Data Deduplication:

In this proposed system, we can detect the duplicate file by using SHA1 algorithm. By using this SHA1, we can create the tags for the uploaded files in the cloud and keys will be generated and maintained in the private cloud. When we upload same file even though file name different, we can identify the duplicate by giving reference id to the newly uploaded file. But, in the cloud only one file will be stored.

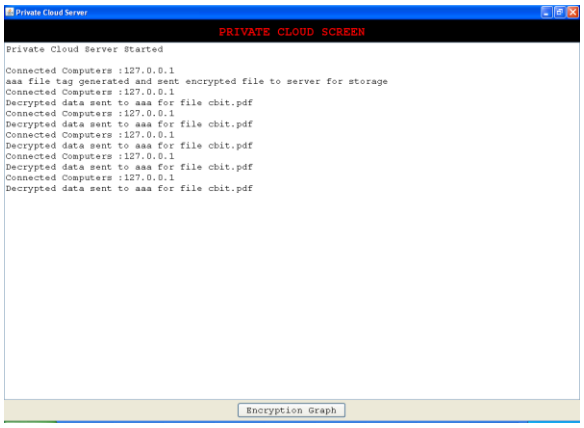
5. RESULT AND DISCUSSIONS

In this experiment we are using private server to deduplicate the cloud files. First, run the private server and here, users can register into the cloud as data owner as well as data users. When the data owner login into the system, he can upload the files into the cloud as well he can share the files to the other users in the cloud. While uploading data he must give the access policies based on the user attributes. After uploading data by data owner, the data users can login into the system and who have the access permission, they only download the files and remaining users cannot download the file. They got message like you don't have access control.

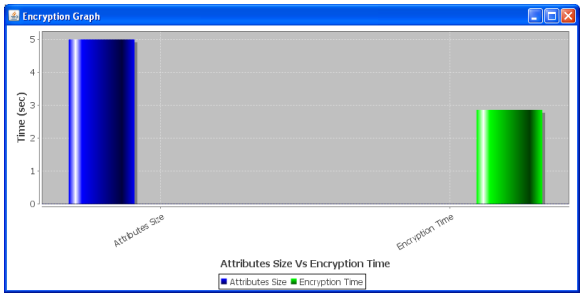
If user is uploading same file again with same or different name, then this application will detect duplicates using Private Cloud Server File Tag and assign reference to old file instead of saving new file.



In above screen we can see for first file reference is none and when same file uploaded then none will replace with old file pointer to avoid duplicates.



Encryption graph can display the number of attributes and processing time of encryption to the whole attributes in the cloud.



6. CONCLUSION

In this paper, we conclude that, we proposed a novel attribute-based totally garage machine assisting at ease deduplication. Our storage device is built under a hybrid cloud architecture, wherein a private cloud manipulates the computation and a public cloud manages the storage. From the experimental results we proved that the proposed system achieved that the share data confidentiality as well as secure data deduplication in the cloud environments.

REFERENCES

[1] 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.

[2] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in Proceedings of the 13th ACM Conference on Computer and Communications Security, CCS 2006, Alexandria, VA, USA, Ioctober 30 - November 3, 2006, ser. Lecture Notes in Computer Science, vol. 5126. Springer, 2006, pp. 89–98.

[3] 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

[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] 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.

[6] 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.

[7] S. Bugiel, S. Nurnberger, 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

[8] 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.

[9] 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

[10] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in 2007 IEEE Symposium on Security and Privacy (S&P 2007), 20-23 May 2007, Oakland, California, USA. IEEE Computer Society, 2007, pp. 321–334.