International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

ACHIEVING EFFICIENT SECURE DEDUPLICATION WITH USER-DEFINED ACCESS CONTROL IN CLOUD

S SHALINI[1], R THANGAPANDIAMMAL[2], Mrs. M.RAJU [3]

[1] [2] M.Sc COMPUTER SCIENCE, [3] ASSISTANT PROFESSOR

DEPARTMENT OF SOFTWARE SYSTEMS

SRI KRISHNA ARTS AND SCIENCE COLLEGE, COIMBATORE

ABSTRACT

By reconfiguring different resources across the Internet, cloud computing offers a cutting-edge method of Offering services. Storage of data is the most significant a popular service on the cloud. Data are regularly stored in the cloud in an encrypted format to protect data owners' privacy. Cloud data deduplication, which is required for processing and storing vast amounts of data on the cloud, is extremely challenging when dealing with encrypted data.Conventional deduplication techniques fail when used to encrypted data. Security issues exist in current encrypted data deduplication methods. They are unable to provide flexible support for data access restriction and revocation. As a result, few of them can be employed in practise without difficulties. Using ownership challenge and proxy re- encryption, we suggest a method for deduplicating Information stored in the cloud that has been encrypted in this study. It integrates data stored in the cloud deduplication along with access control. Utilising in-depth analysis and computer simulations, we grade its performance.The results show how much more effective and efficient the plan is for possiblehands-on use, especially for extensive data deduplication in the cloud.

INTRODUCTION

By rearranging different resources (such storage and processing) and distributing them to consumers in line with their needs, An innovative technique to provide technological services is through cloud computing. By connecting network properties, cloud computing creates a large resource pool. Some of its advantages are scalability, tolerance for faults, flexibility, and pay- per-use. This makes it a potential Model of service right now. Data storage is the most significant and well-liked cloud service. Customers allow a cloud service provider (CSP) access to their private or sensitive data by uploading it to the data centre of the cloud provider. Users of the cloud are encouraged to keep in mind that because attacks and intrusions on sensitive data are likely, CSP cannot be entirely trusted.Furthermore, giving up control of one's own personal data exposes one to significant data security threats, like privacy leaks. The privacy issue grows increasingly significant given how swiftly data mining and other analysis tools are expanding. Therefore, to be able to safeguard user confidentiality and safety of data, it makes appropriate to only upload to the cloud data that is encrypted. However, duplicate data may be uploaded to CSP in encrypted form by the same user or by a number of users, especially when data are shared by many users.Despite the vast amounts of storage that clouds provide, duplicate data severely is a waste of network resources, uses a

much energy, and makes data management difficult. The expansion of various services has made the implementation of efficient resource management techniques even more essential. Therefore, duplication is crucial for large-scale data processing and cloud storage. Duplication has proven to result in significant cost savings. For instance, it can lower storage requirements for file systems and backup programmes by up to 68% and 90-95%, respectively. It is clear that the savings are significant to the economics of cloud company and can be passed on to cloud customers directly or indirectly.Effectively managing encrypted data storage with replication is a practical challenge. However, encrypted data cannot be used with the industrial duplicating methods currently available. Attacks using brute force pose a challenge to replicating currently in use solutions. They are unable to offer both simultaneous flexible data access control and revocation. While dependability, security, and privacy are maintained, performance is often impossible to ensure.

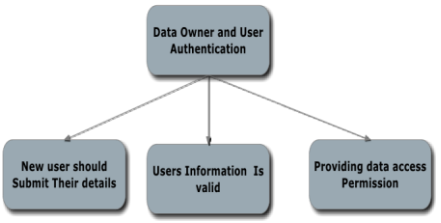
EXISTING SYSTEM

Currently available deduplication methods are susceptible to brute-force assaults. They are unable to permit simultaneous revocation and flexible data access control. Usually, maintaining dependability, security, and privacy does not allow for the guarantee of performance. In order to conserve storage space, providers of cloud storage, including Dropbox, Google Drive, store only one replica of each uploaded file. Deduplication's storage benefits, however, are completely lost if clients generally data encryption. This is due to the usage of multiple encryption keys in order to save the encrypted data as distinct contents. Commercially available systems have difficulty deduplicating encrypted data. Deduplication is a nice illustration of a deduplication technique that is ineffective with encrypted data.

PROPOSED SYSTEM

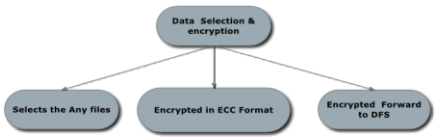
User and Data Owner Authentication

In this module, user and data owner authentications were confirmed. If a user registers as a new user, they must submit their information to the data owner. When a user enters their correct ID and their registration is accepted, their information is acceptable and data access permission is given.



Selection and encryption of data

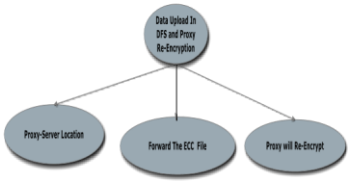
They are unable to offer both simultaneous flexible data access control and revocation. While dependability, security, and privacy are maintained, performance is often impossible to ensure. Companies that provide cloud storage, including Dropbox, Google Drive, Mozy, and others, The data owner selects the files from multiple sources that need to be encrypted with the ECC format. The ECC own key for permission should only be in the possession of the data owner. After that, the file is received by the Distributed File Service (DFS).



Data Upload in Proxy Re-Encryption and DFS

International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

When the ECC file is forwarded to Hadoop's distributed file service.Multiple Proxy-Server Location Information is contained in DFS. Then it forwards the ECC file data to a proxy server, which uses a symmetric encryption technique to re- encrypt the selected file. and keep a server with encrypted data.

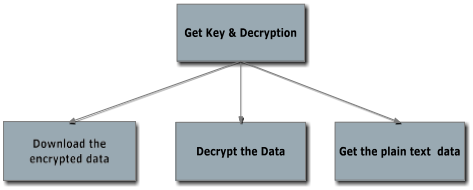


User Data Request and Download

When a user enters this module, their access mode will choose the file and send the request to the data owner. The Request will be forwarded to the DFS by the Data Owner. The user request is forwarded to the proxy server by DFS, which then sends the user's requested files encrypted data along with decrypted key data.

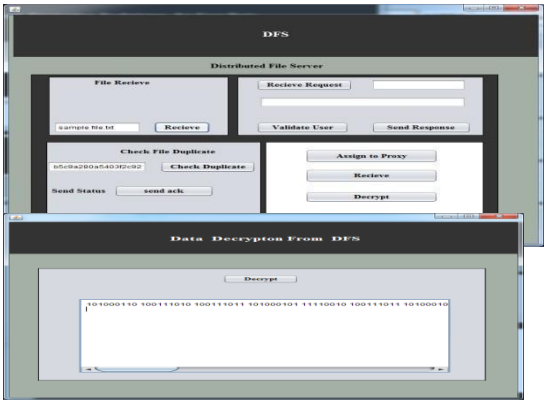
Get Key & Decryption

In this module, the user obtains the ECC file format after employing an encryption key to decrypt encrypted data that was downloaded through a proxy. The user then gets in touch with the data owner to ask for the requisite decryption key in order to obtain the ECC key required to unlock the downloaded file. The user who possesses this key will receive the data in plain text.



RESULTS AND DISCUSSION

Deduplication efficiency findings should be interpreted in terms of space savings and computational overhead. Describe how various deduplication methods strike a balance between effectiveness and performance. Examine how sharing and deduplication affect the balance between security and privacy. Address any worries you may have about shared data assaults and vulnerabilities. The adaptability of your suggested techniques. Think about how well the system functions as data volume and user base grow. Analyse the level of detail in the access control systems. Discuss the advantages of user-specific sharing and potential difficulties in managing access rights for fine-grained access control.Think about the system's usability from the perspectives of the data owner and recipient. Talk about any difficulties users may have securely sharing and gaining access to data.



CONCLUSION

For a cloud storage service to be effective, especially for large data storage, deduplication management of encrypted data is essential and necessary in practise. In the above study, we developed a workable techniqueestablished on ownership challenge and PRE for handling encrypted huge data in the cloud with deduplication. Our technology can update and share data with deduplication even when the data holders are not online..Since only authorised data holders are allowed to obtain Data that is encrypted can be safely accessed if the symmetric keys necessary for data decoding are available. Extensive efficiency evaluations and tests showed that in the context of the security architecture offered, our solution for huge data deduplication works fairly well. The results of our computer models additionally proved the effectiveness of our strategy.

REFERENCES

1. [1] M. Bellare, S. Keelveedhi, and T. Ristenpart, "DupLESS: ServerAided Encryption for Deduplicated Storage," Proceedings of the22nd USENIX Conference on Security, 2013, pp. 179-194.
2. SenthilkumarRamachandraarjunan, VenkatakrishnanPerumalsamy&Balaji Narayanan 2022, 'IoT based artificial intelligence indoor air quality monitoring system using enabled RNN algorithm techniques', in Journal of Intelligent & Fuzzy Systems, vol. 43, no. 3, pp. 2853-2868
3. R.Senthilkumar,Dr.P.Venkatakrishnan,Dr. N.Balaji, Intelligent based novel embedded system based IoT Enabled air pollution monitoring system, ELSEVIER Microprocessors and Microsystems Vol.77, June 2020
4. Dropbox, "A File-Storage and Sharing Service,"http://www.dropbox.com/.
5. Google Drive, http://drive.google.com.
6. Mozy, "Mozy: A File-storage and Sharing Service,"http://mozy.com/.
7. J.R. Douceur, A. Adya, W.J. Bolosky, D. Simon, and M.Theimer, "Reclaiming Space from Duplicate Files in a ServerlessDistributed File System," Proceedings of IEEE InternationalConference on Distributed Computing Systems, 2002, pp. 617 - 624,doi:10.1109/ICDCS.2002.1022312.
8. G. Wallace, F. Douglis, H. Qian, P. Shilane , S. Smaldone, M.Chamness, and W. Hsu, "Characteristics of Backup Workloadsin Production Systems," Proceedings of USENIX Conference onFile and Storage Technologies, 2012, pp. 1-16.Z.O. Wilcox, "Convergent Encryption Reconsidered," 2011,
9. G. Ateniese, K. Fu, M. Green, and S. Hohenberger, "ImprovedProxy Re-Encryption Schemes with Applications to Secure DistributedStorage," ACM Transactions on Information and SystemSecurity, 9(1), 2006, pp. 1-30, doi:10.1145/1127345.1127346.
10. D.T. Meyer and W.J Bolosky, "A Study of Practical Deduplication,"ACM Transactions on Storage, 7(4), pp. 1-20, 2012,doi:10.1145/2078861.2078864.
11. J. Pettitt, "Hash of Plaintext as Key?"http://cypherpunks.venona.com/date/1996/02/msg02013.html.
12. The Freenet Project. Freenet. https://freenetproject.org/.

13. M. Bellare, S. Keelveedhi, and T. Ristenpart, “Message-LockedEncryption and Secure Deduplication,” Proceedings of Cryptology–EUROCRYPT 2013, 2013, pp. 296–312, doi:10.1007/978-3-642-38348-9_18.
14. D. Perttula, B. Warner, and Z. Wilcox-O'Hearn, “Attacks onConvergent Encryption,” http://bit.ly/yQxyvl.
15. C.Y. Liu, X.J. Liu, and L. Wan, “Policy-Based DeduplicationinSecure Cloud Storage,” Proceedings of Trustworthy ComputingServices, 2013, pp. 250–262, doi:10.1007/978-3-642-35795-4_32.
16. P. Puzio, R. Molva, M. Onen, and S. Loureiro, ClouDedup:SecureDeduplication with Encrypted Data for Cloud Storage,”Proceedings of IEEE International Conference onCloud ComputingTechnology and Science, 2013, pp. 363-370,doi:10.1109/CloudCom.2013.54.