International Journal of Science Engineering and Advance Technology

A Data De-duplication Procedure To Protect The Data Security

[1]R.Bindu Madhavi, [2]K.Satya Narayana Murthy
[12]Dept. of CSE, Baba Institute Of Technology & Sciences, Bakkannapalem, Madhurawada,Visakhapatnam ,AP, India

ABSTRACT:

Data de-duplication carry a lot of benefits, security and privacy concerns arise as users' sensitive data are vulnerable to both inside and outside attacks. Traditional encryption, while as long as data confidentiality, is unsuited with data deduplication. Particularly, traditional encryption necessitates different users to encrypt their data with their own keys. Thus, identical data copies of different users will escort to different cipher texts, manufacture deduplication impracticable. To superior guard data security, this paper makes the initial challenge to officially address the dilemma of authorized data deduplication. We implement a prototype of our proposed authorized duplicate check format and manner test bed experiments using our prototype.

KEYWORDS: Deduplication, authorized duplicate check, confidentiality, hybrid cloud.

INTRODUCTION:
To create data management scalable in cloud computing, deduplication has been a well-known method and has paying attention more and more notice lately. Data deduplication is a particular data density technique for get rid of duplicate copies of do again data in storage. The method is used to perk up storage operation and can also be practical to network data transfers to diminish the number of bytes that must be sent. Data deduplication is one of important data density techniques for eradicate duplicate copies of repeating data, and has been broadly used in cloud storage to diminish the amount of storage space and save bandwidth. To look after the discretion of sensitive data while supporting deduplication, the convergent encryption practice has been proposed to encrypt the data before outsourcing.

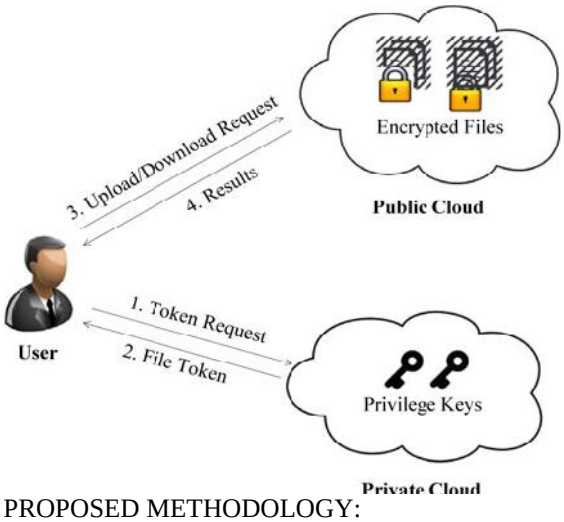
LITERATURE SURVEY:
THE AUTHOR, L. Zhang (ET .AL), AIM this portrays a algorithm which exploits the information which is regular between clients to expand the speed of backups, and reduce the capacity necessities. This algorithm underpins customer end per-client encryption which is essential for private individual information.

THE AUTHOR, W. Lou (ET .AL), AIM we propose Dekey, another development in which clients don't have to deal with any keys all alone however rather safely convey the merged key shares over different servers. Security investigation exhibits that Dekey is secure regarding the definitions indicated in the proposed security display. As a proof of idea, we execute Dekey utilizing the Ramp mystery sharing plan and exhibit that Dekey brings about constrained overhead in sensible situations.

PROBLEM DEFINITION:
As cloud computing turn into widespread, amounting amount of data is being stored in the cloud and communal by users with particular privileges, which classify the access rights of the stored data. One serious challenge of cloud storage services is the running of the ever-increasing volume of data.

PROPOSED APPROACH:
The discrepancy privileges of users are additional careful in duplicate check besides the data itself. We also present more than a few new deduplication constructions supporting authorized duplicate check in hybrid cloud structural design. Security analysis shows that our scheme is protected in terms of the definitions particular in the proposed security model. Deduplication can take place at moreover the file level or the block level. For file-level deduplication, it does away with duplicate copies of the same file. Deduplication can also take place at the block level, which eradicates duplicate blocks of data that come about in non-identical files.

SYSTEM ARCHITECTURE:



PROPOSED METHODOLOGY:
S-CSP.

The S-CSP gives the dataoutsourcing service and stores data on behalf of theusers. To decrease the storage cost, the S-CSP get rid ofthe storage of superfluous data via deduplicationand remain only unique data. In this paper, weassume that S-CSP is always online and has plentifulstorage capacity and calculationauthority.

DATA USERS

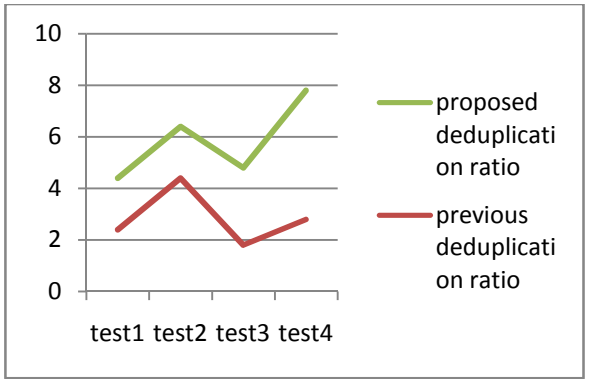
Ina storage system supporting deduplication, the user only uploads exclusive data but does not upload any duplicate data to put away the upload bandwidth, which may be have by the same user or different users. In the authorized deduplication system, each user is concern a set of privileges in the setup of the system.

PRIVATE CLOUD

This is an original entity bring in for make easy user's secure usage of cloud service. specially, because the work out resources at data user/owner side are controlled and the public cloud is not entirely trusted in practice, private cloud is talented to offer data user/owner with an carrying out environment and infrastructure working as an interface between user and the public cloud. The private keys for the privileges are directed by the private cloud, who answers the file token requests from the users.

ALGORITHM:
Authorized duplicate check scheme:
INPUT:F,FT,C,P,H
STEP1:Private cloud maintains a table which contains usersindentity with privileges.
STEP2: Before uploading a file to public cloud data owner performs identification operation send to private cloud server.
STEP3: After passing identification data owner gets file tags.
STEP4: After receiving tag user send to the S-CSP
STEP5: If duplication file is found then
STEP6: User needs to run pow protocol to prove the ownership of file.
STEP7:After passing proof user will get a file pointer.
STEP8: If no duplication is found then
STEP9: Proof will derived to user from S-CSP.
STEP10:User will send proof along with privilege to private cloud server.
STEP11: Private cloud server verifies obtained signature
STEP12: After verification passes user encrypt the file using AES-256 algorithm.
STEP13: User downloads the file by using secretkey.

RESULTS:



The test results are generated by using java language finally shows the performance of de-duplication ratios is effective.

CONCLUSION:

Security study reveals that our methods are secure in terms of insider and outsider attacks particular in the proposed security model. As a verification of thought, we execute a sample of our proposed authorized duplicate check scheme and manner test bed experiments on our prototype. The belief of authorized data deduplication was proposed to look after the data security by with differential privileges of users in the replacement confirm.

REFERENCES:
[1] OpenSSL Project, (1998). [Online]. Available: http://www.openssl.org/

[2] P. Anderson and L. Zhang, "Fast and secure laptop backups withencrypted de-duplication," in Proc. 24th Int. Conf. Large InstallationSyst. Admin., 2010, pp. 29–40.

[3] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Dupless: Serveraidedencryption for deduplicated storage," in Proc. 22nd USENIXConf. Sec. Symp., 2013, pp. 179–194.

[4] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Message-lockedencryption and secure deduplication," in Proc. 32nd Annu. Int.Conf. Theory Appl. Cryptographic Techn., 2013, pp. 296–312.

[5] M. Bellare, C. Namprempre, and G. Neven, "Security proofs foridentity-based identification and signature schemes," J. Cryptol.,vol. 22, no. 1, pp. 1–61, 2009.

[6] M. Bellare and A. Palacio, "Gq and schnorr identification schemesProofs of security against impersonation under active and concurrentattacks," in Proc. 22nd Annu. Int. Cryptol. Conf. Adv. Cryptol.,2002, pp. 162–177.

[7] S. Bugiel, S. Nurnberger, A. Sadeghi, and T. Schneider, "Twinclouds: An architecture for secure cloud computing," in Proc.Workshop Cryptography Security Clouds, 2011, pp. 32–44.

[8] J. R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M.Theimer, "Reclaiming space from duplicate files in a serverlessdistributed file system," in Proc. Int. Conf. Distrib. Comput. Syst.,2002, pp. 617–624.

[9] D. Ferraiolo and R. Kuhn, "Role-based access controls, " in Proc.15th NIST-NCSC Nat. Comput. Security Conf., 1992, pp. 554–563.

[10] GNU Libmicrohttpd, (2012). [Online]. Available: http://www.gnu.org/software/libmicrohttpd/

[11] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Proofs ofownership in remote storage systems," in Proc. ACM Conf. Comput. Commun. Security, 2011, pp. 491–500.

[12] J. Li, X. Chen, M. Li, J. Li, P. Lee, and W. Lou, "Secure deduplicationwith efficient and reliable convergent key management,"in Proc. IEEE Trans. Parallel Distrib. Syst., http://doi.ieeecomputersociety.org/10.1109/TPDS.2013. 284, 2013.

[13] libcurl, (1997). [Online]. Available: http://curl.haxx.se/libcurl/

[14] C. Ng and P. Lee, "Revdedup: A reverse deduplication storagesystem optimized for reads to latest backups," in Proc. 4th Asia-Pacific Workshop Syst., http://doi.acm.org/10.1145/2500727.2500731, Apr. 2013.

[15] W. K. Ng, Y. Wen, and H. Zhu, "Private data deduplication protocolsin cloud storage," in Proc. 27th Annu. ACM Symp. Appl. Comput.,2012, pp. 441–446.

Mrs.R.BinduMadhavi is a student of BABA Institute of Technology and sciences, Visakhapatnam. Presently she is pursuing his M.Tech [CST] from this college and he received his B.Tech from St.Johns collage of Engineering and Technology, affiliated to JNT University, Anantapur in the year 2011. Her area of interest includes Cloud Computing and Object oriented Programming languages, all current trends and techniques in Computer Science.

Mr. K. Satya Narayana Murthy, He has a dedicated teaching experience of three years. He completed B.Tech from GVPCOE, Vizag and M.Tech from Andhra University with distinction and act as a student organizer during M.tech Period. He had begun his Career by joining as a lecturer in BITS vizag. To his ability, he can deal with all the subjects in computer science effectively and got 85% pass percentage in educated subjects. Now he was deeply involved in doing Research, Areas of Research interests includes Data mining, Image Processing, Computer Networks & Cloud Computing. He attended several seminars and workshops and he had conducted various events and organized seminars etc.