ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 1, Issue 2, pp: (89-94), Month: October 2014 – March 2015, Available at: www.paperpublications.org

Secure Deduplication with Efficient and Reliable Dekey Management with the Proof of Ownership

M.Shankari[1], V.Sheela[2], S.Rajesh[3]

[1, 2]Student, Information Technology, Karpaga Vinayaga College of Engineering and Technology, Kanchipuram Dt, Tamilnadu, India
[3]Assistant Professor, Information Technology, Karpaga Vinayaga College of Engineering and Technology, Kanchipuram Dt, Tamilnadu,

Abstract: **De-Duplication improves Storage and bandwidth efficiency is incompatible with traditional encryption. In traditional model encryption requires different users to encrypt their own data with their own master key, thus identical data copies of different users will lead to different cipher texts, making de-duplication impossible. Each such copy can be defined based on different granularities: it may refer to either a whole file (i.e., file level deduplication), or data block (i.e., block-level deduplication). To applying deduplication to user data to save maintenance cost in cloud. Apart from normal encryption and decryption process we have proposed Master key concept with DeKey concept. For Encryption and Decryption we have used Triple Data Encryption Standard Algorithm where the plain text is encrypted triple times with the key so that the data is secure and reliable from hackers. We reduced the cost and time in uploading and downloading with storage space.**

Keywords: **De-Duplication, Encryption and Decryption.**

1. INTRODUCTION

The advent of cloud storage motivates enterprises and organizations to outsource data storage to third party cloud providers, as evidenced by many real-life case studies. One critical challenge of today's cloud storage services is the management of the ever-increasing volume of data. According to the analysis report, the volume of data in the wild is expected to reach 40 trillion gigabytes in 2020. To make data management scalable, deduplication has been a well-known technique to reduce storage space and upload bandwidth in cloud storage. Instead of keeping multiple data copies with the same content deduplication eliminates redundant data by keeping only one physical copy and referring other redundant data to that copy. Each such copy can be defined based on different granularities: it may refer to either a whole file (i.e., filelevel deduplication), or a more fine-grained fixed size or variable-size data block (i.e., block-level deduplication).Today's commercial cloud storage services, such as Drop box, Mozy, and Memo pal, have been applying deduplication to user data to save maintenance cost .From a user's perspective, data outsourcing raises security and privacy concerns. We must trust third-party cloud providers to properly enforce confidentiality, integrity checking, and access control mechanisms against any insider and outsider attacks. However, deduplication, while improving storage and bandwidth efficiency, is incompatible with traditional encryption. Specifically, traditional encryption requires different users to encrypt their data with their own keys. Thus, identical data copies of different users will lead to different cipher texts, making deduplication impossible.

2. EXISTING SYSTEM

In existing system data De-duplication is not scalable. It suffers in two issues first; it is inefficient, as it will generate an enormous number of keys with the increasing number of users. Specifically, each user must associate an encrypted

ISSN 2350-1022

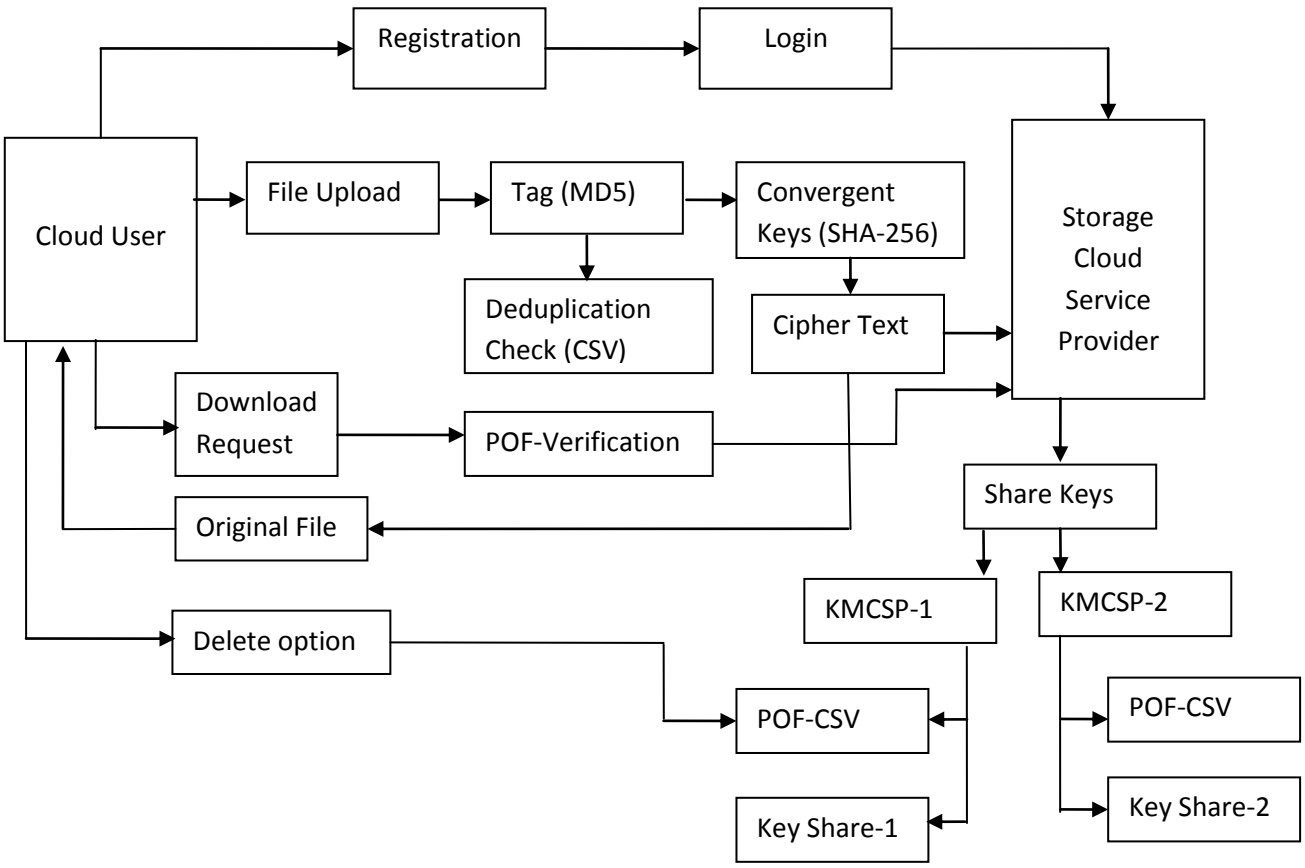
International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 1, Issue 2, pp: (89-94), Month: October 2014 – March 2015, Available at: www.paperpublications.org

convergent key with each block of its outsourced encrypted data copies, so as to later restore the data copies. Although different users may share the same data copies, they must have their own set of convergent keys so that no other users can access their files. As a result, the number of convergent keys being introduced linearly scales with the number of blocks being stored and the number of users. Second, the baseline approach is unreliable, as it requires each user to dedicatedly protect his own master key. If the master key is accidentally lost, then the user data cannot be recovered; if it is compromised by attackers, then the user data will be leaked. Cost increases to the storage of content as well as for the keys storage. Increase bandwidth with upload time. Security lacks with only Master Key Concept.

3. PROPOSED SYSTEM

DeKey concept do not relay on Master Key Concept. We have shown the concept of deduplication effectively and security is achieved by means of Proof of Ownership of the file. We outsource the convergent keys to third party key Management server securely. DeKey supports both file-level and block level deduplications. Key Management overhead is avoided and provides fault tolerance guarantees for key management, while preserving the required security properties of secure deduplication. Instead of using normal encryption and decryption we use Triple DES Technique as the plain text is encrypted triple times with the convergent key so that our data will be secured. Scalability increases as DeKey achieved efficiently. Cost efficiency is achieved as multiple users of same date is just referred and not newly added. Deleting content of shared file of different user will allow deleting only convergent keys references not content stored in server.

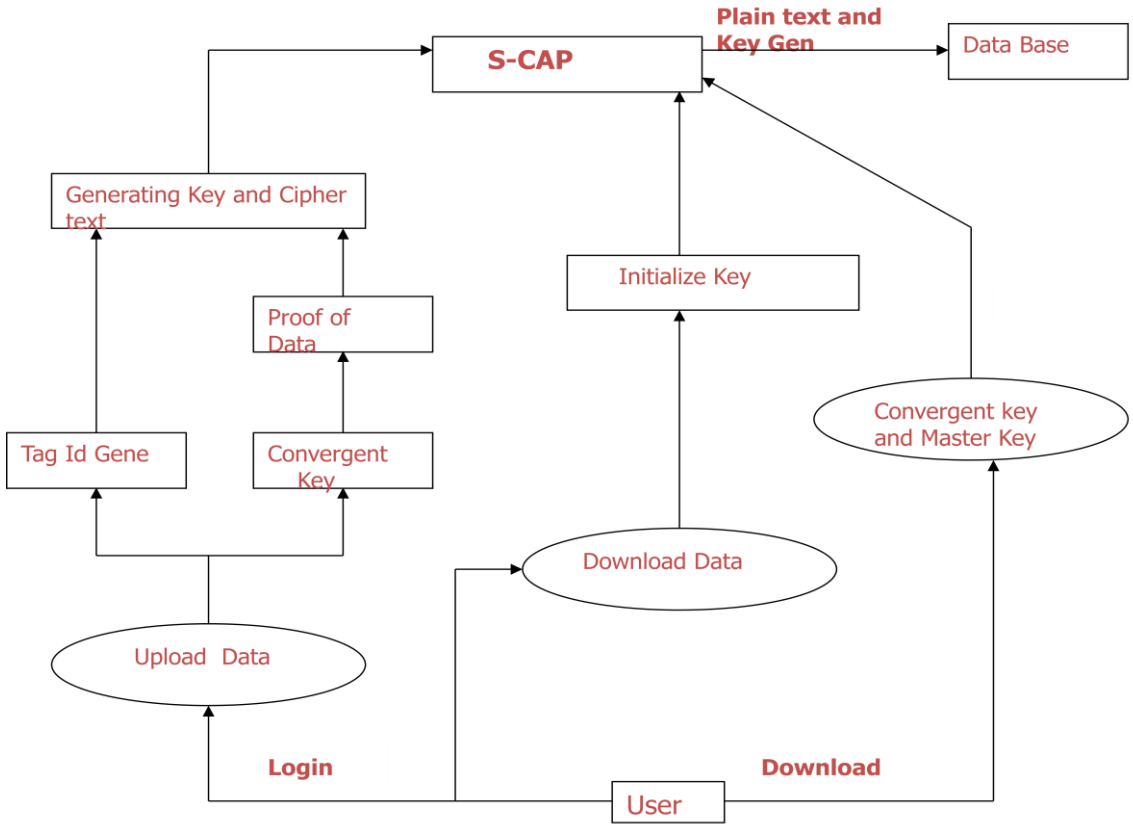
ARCHITECTURE:



ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 1, Issue 2, pp: (89-94), Month: October 2014 – March 2015, Available at: www.paperpublications.org

4. SYSTEM MODEL



MODULES:

1. Mastering file to cloud service provider.

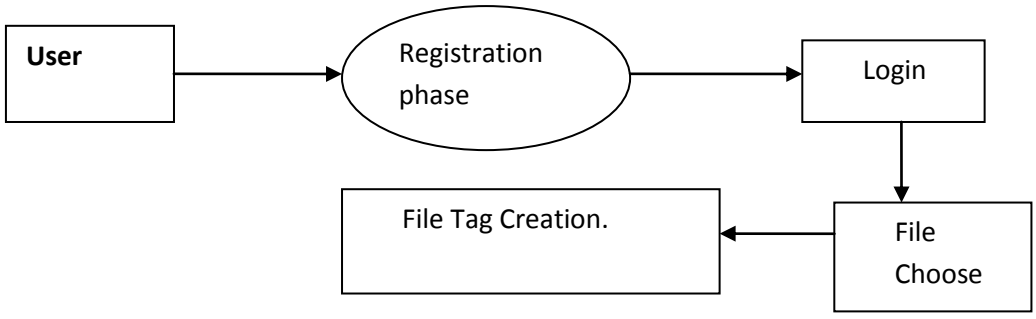
2. Chucking the file chosen.

3. DeKey based encryption.

4. Hash value based decryption.

4.1 MASTERING FILE TO CLOUD SERVICE PROVIDER:

A user is an entity who wants to outsource data storage to the storage cloud service provider (S-CSP) and access the data later. User registers to the cloud server with necessary information and login cloud page for uploading the file. User chooses the file and uploads to server where the server store the file in rapid storage system and file level de-duplication is checked. We tag the file 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.

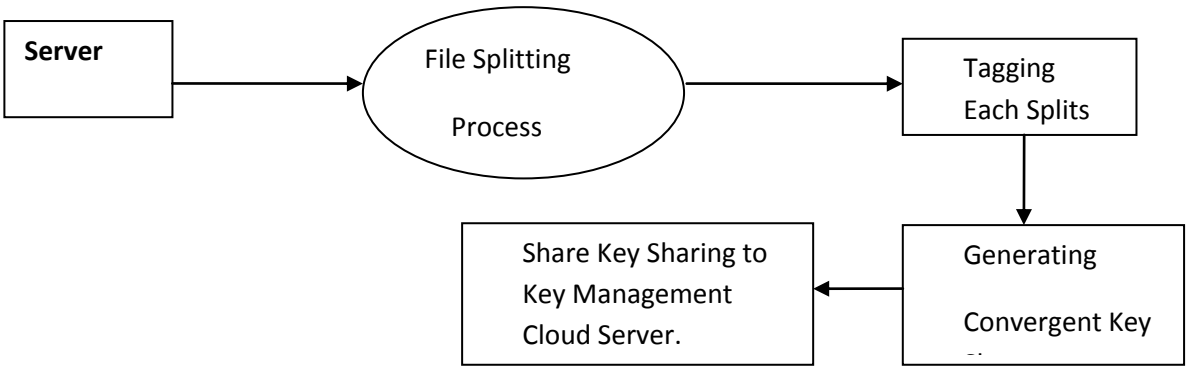


ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 1, Issue 2, pp: (89-94), Month: October 2014 – March 2015, Available at: www.paperpublications.org

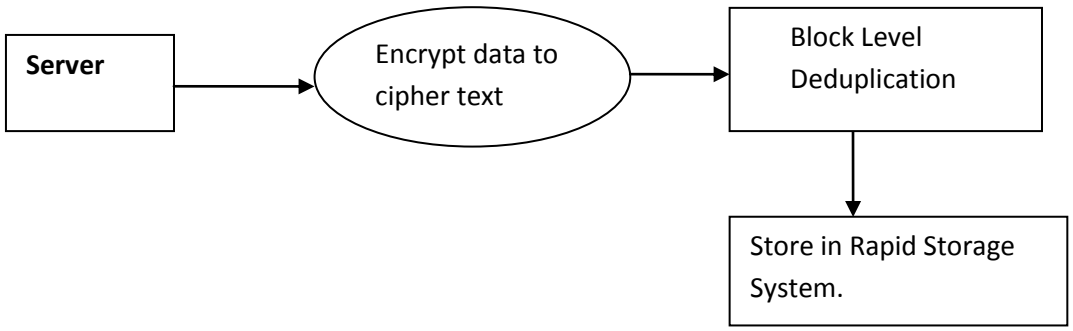
4.2 CHUCKING THE FILE CHOSEN:

Chunking the file chosen of fixed size and generating tags for each blocks chunked. After that generate convergent keys for each blocks split to verify block level deduplication. Here we provide filename and password for file authorization in future. Encrypt the blocks by Triple Data Encryption Standard (3DES) algorithm. Here the plain text is encoded triple times with convergent key and so the while decoding the original content it also need the same key to decode again by triple times. Finally the original content is encrypted as cipher text and stored in Storage Cloud Service Provider (S-CSP) file storage system.



4.3 DEKEY BASED ENCRYPTION:

After encryption the convergent keys are securely shared with cloud service provider to Key Management Cloud Service Provider (KMCSP). Key management server checks duplicate copies of convergent keys in KMCSP. Key Management Server maintains Comma Separated Values (CSV) file to check proof of verification and store keys secure. The different users who share the common keys are referred by their own ownership. User request for deletion definitely need to prove proof of ownership to delete own contents.

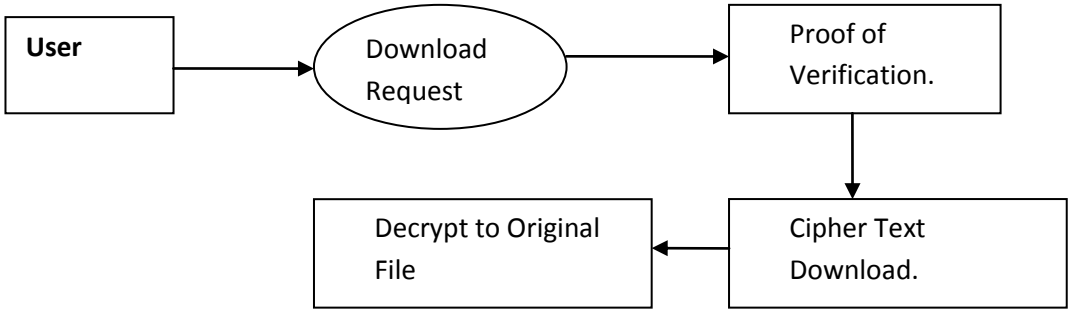


4.4 HASH VALUE BASED DECRYPTION:

The final model where the user request for the downloading their own document which they have been upload and stored in cloud server. This download request needs proper ownership verification of the document here we create the ownership by unique tag generated by MD5 algorithm and verifies existing tag of user. After verification the original content is decrypted by requesting the cloud server where cloud server request key management server for keys to decrypt and finally the original content is received by the user. The delete request will delete only the reference of the content shared by common users and not the whole content.

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 1, Issue 2, pp: (89-94), Month: October 2014 – March 2015, Available at: www.paperpublications.org



5. RELATED WORDS

M.Bellare,S.keelveedhi[1], new approach to multi-level secure multicast, one that uses the secure lock encryption scheme based on the mathematics of the Chinese Remainder Theorem .D.Harnik,B.Pinkas[2], With the growing data size of cloud computing, a reduction in data volumes could help providers reducing the costs of running large storage system and saving energy consumption . So data deduplication techniques have been brought to improve storage efficiency in cloud storages dynamic deduplication scheme for cloud storage, which aiming to improve storage efficiency and maintaining redundancy for fault tolerance.S.Kamara,K.Lauter[3], Due to vast availability of resources and numerous tasks security is an important concern in cloud environment .M.Li [4],new architectures, algorithms, data collection and evaluation methods.D.Meister.A.Brikmann[5] , the tenant performs customization separately for different business applications, which brings about many repeated custom operations and much duplication of custom metadata.M.W.Storer,K.Greenn[6], Proof of Ownership (POW) improves storage efficiency by securely removing unnecessarily duplicated data on the storage server. However, trivial combination of the two techniques, in order to achieve both data integrity and storage efficiency, results in non-trivial duplication of metadata (i.e., authentication tags), which contradicts the objectives of POW. Recent attempts to this problem introduce tremendous computational and communication costs and have also been proven not secure. Y.Tang, P.P.Lee, J.C.Lui [7], the system's two key modules, at the Brammer application, and at another Micross installation.G.Wallace, F.Douglis [8], a secure overlay cloud storage system that achieves fine-grained, policy-based access control and file assured deletion. It associates outsourced files with file access policies, and assuredly deletes files to make them unrecoverable to anyone upon revocations of file access policies .Q. Wang, C.Wang [9],. Data integrity auditing and storage deduplication are achieved simultaneously. Our proposed scheme is also characterized by constant real-time communication and computational cost on the user side. Cloud computing based storage services have rapidly spread in the market due to their promising capabilities and features. However, the security challenge of outsourcing sensitive data for sharing on the cloud which is not fully controlled by the data owners is still open. W. Wang, B.Bargava [10], we design a cipher text-policy attribute-based encryption (ABE) scheme and a proxy re-encryption scheme. Based on them, we further propose a secure, efficient and fine-grained data Access Control mechanism for P2P storage Cloud named ACPC.

6. CONCLUSION

We propose DeKey, an efficient and reliable convergent key management scheme for secure deduplication. DeKey applies deduplication among convergent keys and distributes convergent key shares across multiple key servers, while preserving semantic security of convergent keys and confidentiality of outsourced data. We implement DeKey using the Proof of ownership scheme and demonstrate that it incurs small encoding/decoding overhead compared to the network transmission overhead in the regular upload/download operations.

REFERENCES

[1] M. Bellare, S. Keelveedhi, and T. Ristenpart, ‘‘Message-Locked Encryption and Secure Deduplications,’’ in Proc. IACR Cryptology ePrint Archive, 2012, pp. 296-3122012:631.

[2] D. Harnik, B. Pinkas, and A. Shulman-Peleg, ‘‘Side Channels in Cloud Services: Deduplications in Cloud Storage,’’ IEEE Security Privacy, vol. 8, no. 6, pp. 40-47, Nov./Dec. 2010.

[3] S. Kamara and K. Lauter, ‘‘Cryptographic Cloud Storage,’’ in Proc. Financial Cryptography: Workshop Real-Life Cryptograph. Protocols Standardization, 2010, pp. 136-149.

[4] M. Li, ‘‘On the Confidentiality of Information Dispersal Algorithms and their Erasure Codes,’’ in Proc. CoRR, 2012, pp. 1-4abs/1206.4123.

[5] D. Meister and A. Brinkmann, ‘‘Multi-Level Comparison of Data Deduplications in a Backup Scenario,’’ in Proc. SYSTOR, 2009, pp. 1-12.

[6] M.W. Storer, K. Greenan, D.D.E. Long, and E.L. Miller, ‘‘Secure Data Deduplications,’’ in Proc. StorageSS, 2008, pp. 1-10.

[7] 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.

[8] G. Wallace, F. Douglis, H. Qian, P. Shilane, S. Smaldone, M. Chamness, and W. Hsu, ‘‘Characteristics of Backup Workloads in Production Systems,’’ in Proc. 10th USENIX Conf. FAST, 2012, pp. 1-16.

[9] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, ‘‘Enabling Public Audit ability and Data Dynamics for Storage Security in Cloud Computing,’’ IEEE Trans. Parallel Distrib. Syst., vol. 22, no. 5, pp. 847-859, May 2011.

[10] W. Wang, Z. Li, R. Owens, and B. Bhargava, ‘‘Secure and Efficient Access to Outsourced Data,’’ in Proc. ACM CCSW, Nov. 2009, pp. 55-66.