ENHANCING THE SECURE CLOUD STORAGE IN CLOUD ENVIRONMENTS

P. Santosh Babu[1], G. Minni[2], Syeed Yasin[3]
[1]M. Tech Student, Department of IT, [2]Asst. Professor, [3]Associate. Professor, Department of CSE,
Nimra College of Engineering & Technology, Vijayawada, Andra Pradesh, India.

Abstract: One main challenge of today's cloud storage services is that the management of the ever increasing amount of data. to create data management protractible de-duplication has been a standard technique to compress hold and transfer bandwidth in cloud storage rather than keeping multiple information copies with an identical content, de-duplication eliminates redundant information by keeping just one physical copy and referring alternative redundant data to that copy. Currently every day the foremost arising challenge is to perform secure de-duplication in cloud storage. though convergent cryptography has been extensively maintain for secure de-duplication, a important issue of making convergent cryptography wise is to with efficiency and faithfully to manage a large session of convergent keys. we tend to initial introduce a normal approach therein every user holds associate self master key for encrypting the convergent keys and outsourcing them to the cloud. However, such a standard key management scheme generates large variety of keys with the increasing number of users and desires users to reserved for a particular use to defend the master keys that is inefficient an unreliable. For that purpose we tend to are about to formally address the matter of achieving efficient and reliable key management in secure de-duplication. Organizations usually use de-duplication in backup and failure recovery applications. During this paper we tend to try to approve de-duplication check similarly as file compression, combine with convergent coding for providing security to sensitive data using hybrid cloud.

Index Terms: Hybrid Cloud, Convergent Encryption, De-duplication, Proof of Ownership;

I. INTRODUCTION

Cloud computing is getting more and more widespread because it can give low-priced and on demand use of huge storage and method resources. As the volume of data grows, in addition increasing is the Total value of ownership that includes storage infrastructure value, management worth and human administration price. therefore in cloud storage systems, reducing the quantity of information that want to be keep, transferred and managed becomes a crucial. As a result, information De-duplication is an necessary and standard cost-saving feature for cloud storage .The term information de-duplication refers to techniques that store just one copy of unnecessary information, and provides links to that duplicate instead of storing different actual copies of this data. With the transition of services from tape to disk, information de-duplication has become a key component in the backup method. By storing and forwarding only one copy of duplicate information, de-duplication offers savings of every disk space and network information measure. De-duplication could be a one of the vital, technique to reduce storage space and transfer metric and has been accustomed build information management ascendible .As an different of keeping varied information copies with the identical content, de-duplication eliminates unused information by keeping solely one physical copy and referring different unused information to that copy. There are two sorts of de-duplication check one is file-level de-duplication and another is block-level de-duplication. Among that file-level de-duplication introduce the complete file where as block-level de-duplication refers to the mounted or variable size information block. To make de-duplication secure we have to use sure security mechanism like cryptography. Traditional cryptography wants whole different users to cipher their knowledge with their own keys, therefore identical information copies of totally different users can cause different cipher text and for this reason de-duplication is incompatible with ancient cryptography. Convergent cryptography provides a potential possibility to implement information confidentiality whereas produce the de-duplication. convergent cryptography, a cryptosystem that produces in distinguishable cipher text files from an identical plaintext files, regardless of their cryptography keys it encrypts or decrypts an information with a cryptography key, that's derived by computing the encoded hash worth of the content of the information copy itself. Once key generation and information cryptography, users retain the keys and send the encoded-text to the cloud. Since cryptography is settled, identical information copies can generate constant convergent key and also the same encoded-text. This permits the cloud to perform de-duplication on the encoded text. The encoded-texts can only be decrypted by the corresponding information house owners with their merging keys. To stop uncertified access, a secure proof of ownership (POW) protocol is additionally required to produce t he proof that the user therefore owns an identical file once a duplicate is found. Once the proof, subsequent users with an identical file are assign a pointer from the server while not having to upload an equivalent file. A user can download the encoded file with the pointer from the server, which can only be decrypted by the corresponding information owners with their convergent keys. Thus, convergent secret writing permits the cloud to perform de-duplication on the cipher texts and additionally the proof of owner-ship stop the uncertified user to access the file by using convergent cryptography technique we will observe duplicate files likewise file compression. Through our implementation we will observe the duplicate files still as we will increase the storage space of cloud.

II. RELATED WORK

Symmetric cryptography additionally referred to as typical cryptography or single key encoding was the only quite cryptography in use before the event of public-key encoding. The symmetric cryptography theme has five ingredients like: 1. Plaintext: this is the initial intelligible message or information that is fed to the algorithm as input. 2. Encryption algorithm: The encoding rule performs varied sub interchange and permutations on the plaintext. 3. Secret Key: the key key's additionally input to the cryptography rule. The precise substitutions and permutations performed depend upon the key used, and the algorithm will produce a novel output counting on the specific key getting used at the time. 4. Cipher text: this is often the disordered message created as output. It depends on the readable-text and therefore the key. The cipher text is associate apparently random stream of data, because it stands, is in comprehensible. 5. Coding Algorithm: this is often essentially the cryptography rule run in reverse. It takes the cipher text and the secret key and produces the original plaintext.

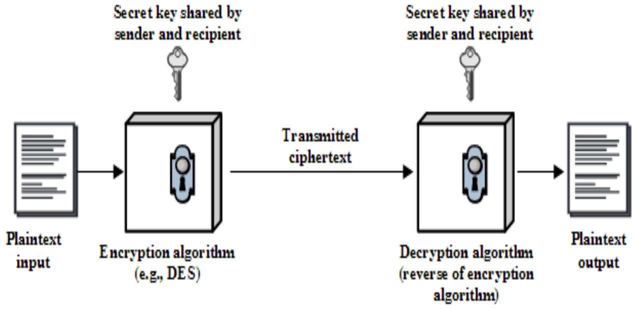

Secret key shared by sender and recipient
Secret key shared by sender and recipient
Transmitted ciphertext
Plaintext input
Encryption algorithm (e.g., DES)
Decryption algorithm (reverse of encryption algorithm)
Plaintext output

Figure 1: Architecture of Existing System

We assume it is impractical to decipher a message on the premise of the cipher text and data of the encryption/decryption rule. In different words, we tend to do not need to keep the rule secret we wish to keep only the key secret. Sender and also the receiver should have obtained copies of the key in a secure fashion and should keep the key secure. If someone can discover the key and knows the rule, all communications using this secret key is readable. We are able to describe however we will use a public-key cryptosystem for a secure key exchange later in this instruct.

III. FRAME WORK

As the name suggests a Hybrid cloud is also a mixture of each a public and private cloud as an example like organization or a corporation or a company may prefer to place their operative settings in a very public cloud whereas the event environment is additionally placed during a very personal cloud. additionally several organizations like better to run their sales and marketing operations during a public cloud, whereas keeping their money operations among a private cloud another choice is two run twin systems. Private cloud we tend to area unit storing encoded keys. Non-public cloud is high flexibility and high worth and high secure than compare to public cloud. Public Cloud we tend to are storing the encoded text or unclear text. Public cloud is low flexibility and low value and less secure than compare to non-public cloud. The Secure-Cloud Storage suppliers is an entity that provides the secure information storage service for the users. In the de-duplication system, once users hold and store equal information, the S-CSP will only store one copy of those files and retain only distinctive data. A de-duplication technique, on the further offer to reduce the storage cost at the server side and save the transfer bandwidth at the user aspect. For backup and confidentiality of information storage, we tend to consider a gathering of Secure-cloud storage providers. Information De-duplication involves finding and removing of duplicate information without considering its fidelity .Here the goal is to store lots of information with less bandwidth and price.

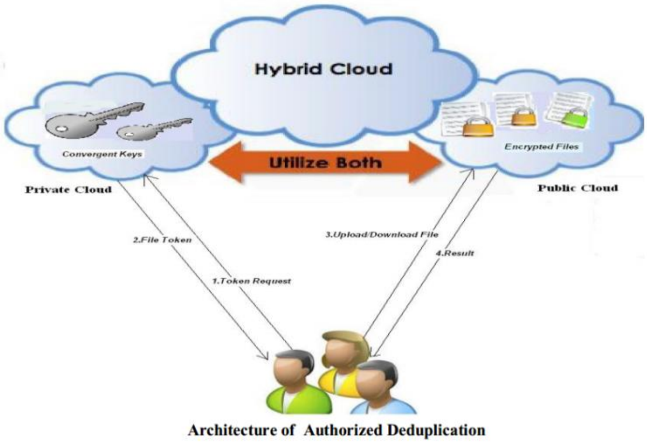


Figure 2: Architecture of Hybrid Cloud

Convergent coding provides information confidentiality in de-duplication. A user (or data owner) derives a merging key from every original information copy and encrypts the information copy with the convergent key. The basic set up of convergent encoding (CE) is to derive the cryptography key from the hash of the readable text. the sole implementation of convergent cryptography are going to be defined as follows: Alice derives the cryptography key from her file M such that K = H(M), where H is a cryptographic hash perform he will encode the message with this key, hence: C = E(K,M) = E(H(M), where E is a block encoded. By applying this technique, two users with two identical plain texts can acquire two identical encoded texts since the cryptography key is the same; therefore the cloud storage :ownership, may be able to perform de-duplication on such as encoded texts. Moreover, coding keys are generated, maintained and protected by users. As the encoded key is deterministically generated from the plain text, users do not need to interact with each other for agreement on the key to cipher a given plaintext. Therefore the convergent cryptography appear to be a wise candidate for the adoption of cryptography and de-duplication in the cloud storage domain. In addition, the user verify a tag for the information copy, specified the tag are used to discover duplicates.

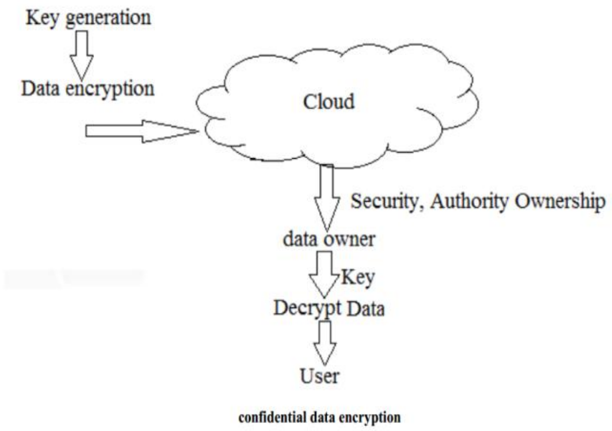


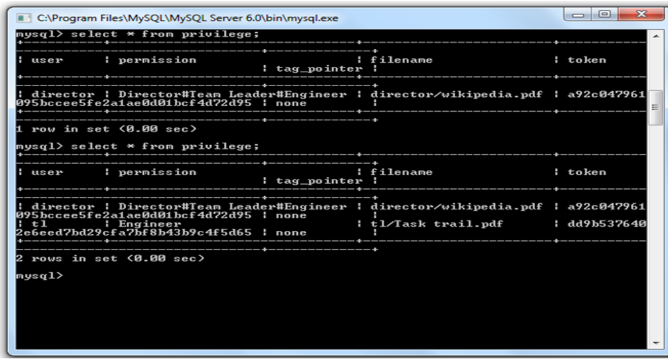
Figure 3: Convergent Encryption Mechanism

A convergent cryptography theme are defined with four primitive functions: Key Gen(M) → K is the key generation algorithm that maps a data copy M to a merging key K Encrypt(K,M) → C is the symmetric cryptography formula that takes each the encoded key K and so the data copy M as inputs then outputs a encoded text Decrypt(K,C)→M is the decoding algorithm that takes each the cipher text C and so the merging key K as inputs then outputs the initial data copy M Tag-Gen(M) →T(M) is the tag generation algorithm that maps the original data copy M and outputs a tag T(M). we tend to enable Tag-Gen to generate a tag from the corresponding cipher text by using T(M)=Tag-Gen(C), where C=Encrypt(K,M). The notion of proof of rights allows users to prove their ownership of data copies to the storage server.

IV. EXPERIMENTAL RESULTS

In our experiments, any number of users can registered and login into the system. Who are authorized users they can upload the files into the cloud. Any user can give the access permission to other authorized users after upload the files. These uploaded files are stored in public cloud and keys are stored in private cloud. If any duplicate files are available in public cloud, then that file cannot uploaded in the cloud and to that particular file tag pointer will be assigned to the user. But that file can be downloaded by data owner as well as data users.

Below image shows that the tag pointers and access permission can view in the database;



```
mysql> select * from privilege;
+----------+------------------------------+------------------------+-----------
-----------------------------+-----------------+
| user     | permission                   | filename               | token
                             | tag_pointer |
+----------+------------------------------+------------------------+-----------
-----------------------------+-----------------+
| director | Director#Team Leader#Engineer | director/wikipedia.pdf | a92c047961
095bccee5fe2a1ae0d01bcf4d72d95 | none        |
+----------+------------------------------+------------------------+-----------
-----------------------------+-----------------+
1 row in set (0.00 sec)

mysql> select * from privilege;
+----------+------------------------------+------------------------+-----------
-----------------------------+-----------------+
| user     | permission                   | filename               | token
                             | tag_pointer |
+----------+------------------------------+------------------------+-----------
-----------------------------+-----------------+
| director | Director#Team Leader#Engineer | director/wikipedia.pdf | a92c047961
095bccee5fe2a1ae0d01bcf4d72d95 | none        |
| tl       | Engineer                     | tl/Task trail.pdf      | dd9b537640
2e6eed7bd29cfa7bf8b43b9c4f5d65 | none        |
+----------+------------------------------+------------------------+-----------
-----------------------------+-----------------+
2 rows in set (0.00 sec)

mysql>
```

Through our implementation we can store the data into the cloud in public cloud and encoded keys are stored in private cloud also detect the duplicate files as well as we can increase the storage space of cloud and also we can decrease the network bandwidth and cost.

V. CONCLUSION

Cloud computing has reached a majority, that leads it into a productive section. This suggests that most of the main issues with cloud computing are addressed to a degree that clouds have become interesting for full industrial exploitation. This however can not mean that all the problems listed higher than have truly been solved , only that the according risks can be tolerated to a sure measure. Cloud computing is therefore still as much a research topic, because it may be a market providing. For higher confidentiality and security in cloud computing we tend to have projected new de-duplication constructions supporting approved duplicate check additionally as file compression in hybrid cloud style, in this the duplicate-check tokens of files are generated by private cloud server with non-public keys. Projected system includes proof of data owner so it will facilitate to implement higher security problems in cloud computing.

REFERENCES

[1] ANDERSON, P., ANDZHANG, L. Fast and secure laptop backups with encrypted de-duplication. In Proc. of USENIX LISA (2010).

[2] BELLARE,M., BOLDYREVA, A.,ANDO'NEILL, A. Deter-ministic and efficiently searchable encryption. InCRYPTO 2007(Santa Barbara, CA, USA, Aug. 1923,2007), A. Menezes, Ed.,vol. 4622 ofLNC, Springer, Berlin, Germany, pp. 535–552.

[3] BELLARE,M.,KEELVEEDHI,S.,ANDRISTENPART,T. Message-locked encryption and secure de-duplication. In EU-ROCRYPT 2013, to appear. Cryptology ePrint Archive, Report 2012/631, November 2012.

[4] Pasqualo Puzio, Refik Molva, MelekOnen,"CloudDedup: Secure Deduplication with Encrypted Datafor Cloud Storage", SecludIT and EURECOM,France.

[5] W. K. Ng, Y. Wen, and H. Zhu, "Private data deduplication proto cols in cloud storage," in Proc. 27th Annu. ACM Symp. Appl. Com-put., 2012, pp. 441–44.

[6] M. W. Storer, K. Greenan, D. D. E. Long, and E. L. Miller, "Secure data deduplication," in Proc. 4th ACM Int. Workshop Storage Security Survivability, 2008, pp. 1–10.

[7] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Proofs of ownership in remote storage systems," in Proc. ACM Conf. Com-put. Commun. Security, 2011.

[8] M. Shyamala Devi, V.Vimal Khanna,Naveen Balaji "Enhanced Dynamic WholeFile De-Duplication(DWFD) for Space Optimization in PrivateCloudStorage Backup",IACSITAugust,2014.

[9] Weak Leakage-Resilient Client–Side deduplication of Encrypted Data in Cloud Storage" Institute for

Info Comm Research,Singapore,2013

[10] BELLARE, M., KEELVEEDHI, S., ANDRISTENPART, T.Message-locked encryption and secure deduplication. InEU-OCRYPT 2013, toappear. Cryptology ePrint Archive, Report2012/631, November 2012.

[11] OpenSSL Project. http://www.openssl.org/.

[12] GNU Libmicrohttpd. http://www.gnu.org/software/libmicrohttpd/.

[13] 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

[14] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman,“Role-based access control models,”IEEE Comput., vol. 29, no. 2,pp. 38–47, Feb. 1996

[15] J. Yuan and S. Yu, “Secure and constant cost public cloud storageauditing with deduplication,” IACR Cryptology ePrint Archive, 2013:149, 2013.