An Enhanced Multi Layered Cryptosystem Based Secure and Authorized Deduplication Model in Cloud Storage System

Ch.Venkatesh
Computer Science and Technology
VR Siddhartha Engineering College
Vijayawada, kanur, India
venkateshchallapalli1@gmail.com

S. Kranthi
Information Technology
VR Siddhartha Engineering College
Vijayawada, kanur, India
kranthisri41@gmail.com

Abstract: - As the cloud computing science develops for the duration of the last decade, outsourcing knowledge to cloud service for storage becomes an attractive pattern, it's nothing but remote gaining access to on heavy knowledge renovation and management. The content material shared by way of extraordinary privileges customers; arguably it raises security issues to guard the authorized knowledge whilst supporting data deduplication. Data Deduplication is one in all principal data compression strategies for taking away duplicate copies of repeating information, and has been commonly utilized in cloud storage to diminish the quantity of cupboard space and keep bandwidth. As more company and private users outsource their data to cloud storage providers, data breach incidents make end-to-end encryption. The convergent encryption technique proposed to encrypt the data earlier than outsourcing which provides a protocol that decrypt the ciphertext to supply data from the private to public cloud. The Advance encryption standard (AES) enhances the token generation for each and every user, Hybrid cloud appoach is introduced. Data Deduplication is one in every of foremost information compression procedures for eliminating duplicate copies of repeating data, and has been broadly utilized in cloud storage to minimize the amount of storage space and bandwidth.

Keywords— Data Deduplication, Convergent Encryption, AES, Hybrid.

I. INTRODUCTION

The Cloud computing supplies seemingly limitless virtualized assets to customers as services throughout the whole web. As cloud computing turns into ordinary, an increasing quantity of data is being saved within the cloud and shared via users with precise privileges, which define the access rights of the stored knowledge. One relevant assignment of cloud storage offerings is the management of the ever-growing quantity of information. To make information administration scalable in cloud computing, Deduplication has been a well-known method and has attracted more and more attention not too long ago.

Knowledge deduplication is a specialised information compression process for removing duplicate copies of repeating data in storage. The procedure is used to enhance storage utilization and can also be applied to network data transfers to minimize the number of bytes that have to be despatched. Alternatively of retaining multiple information copies with the equal content, deduplication eliminates redundant information by keeping only one physical copy and referring other redundant data to that copy. [2] Deduplication can take position at either the file level or the block degree. For file degree deduplication, it eliminates replica copies of the identical file. Deduplication can also take position on the block level, which eliminates replica blocks of information that occur in non-same files.

Despite the fact that knowledge deduplication brings a number of benefits, safety and privateness concerns arise as users sensitive information are inclined to both insider and outsider assaults. Convergent encryption has been greater to put in force data confidentiality while making deduplication possible. [5] It encrypts/ decrypts a data replica with a convergent encryption is bought via computing the cryptographic hash value of the content of the data copy. After key iteration and information encryption, customers keep the keys and ship the ciphertext to the cloud. The important thing cannot be view or shared or stolen by the opposite person which makes a lot stronger in security and the deduplication is finished in not best file smart, deduplication is finished in content wise and dimension intelligent within the file size. So there is no risk of repeated information in database of the given precise privilege. Advanced Encryption Standard (AES) is offered for encrypt/decrypt the file, Hybrid cloud appoarch is carried out the place combination of both private and public cloud concerned. Drivehq is the public cloud the places records are get in encrypted format.

II. PROBLEM STATEMENT

The problem rises in mainly security, wastage of storage space and bandwidth consumption. The Advance encryption Standard (AES) enhances the ciphertext into plain text by encrpyt/decrpyt format. Traditional convergent encryption Technique in providing unique ID, Indexing deduplicate file, better security, reduce storage cost and save bandwidth for better performance.

Efficiently solving the problem of deduplication with differential privileges/users in cloud computing. Hybrid cloud architecture i.e Database as a service provided where both private and public cloud is used for solving our problem. As more corporate and private users outsource their data to cloud storage providers, data breach incidents make end-to-end encryption. Adavance encryption Standard is used for better security where same key cannot share by different users.

Enhancing design an encryption scheme that guarantees semantic security for unpopular data and provides weaker security and better storage and bandwidth benefits for popular data to avoid duplicate copies. The storage space will wasted for unnessary data while takes to redundancy and loss consistency. So Deduplication concept duplicate checks the content inside the file exists and even file size with respective same filename.

III. STUDY OF PROCEDURE

It involves different phases that the complete study of procedure that are enhances.

A. Symmetric Encryption

It uses a common secret key K to encrypt and Decrypt the information. This Scheme consists of three primitive functions.

1) Key generation->K is key generation algorithm that generates k using symmetric parameter 1.

2) Encryption symmetric (K,M)->C is symmetric encryption algorithm that takes secret K and Message M and also outputs the ciphertext C.

3) Decryption symmetric (K,C)->M is the symmetric decryption algorithm that takes the secret K and ciphertext C and then outputs the original message M.

B. Convergent Encryption

1) Key generation (M)->K is key generation algorithm that maps a data copy M to a convergent key K.

2) Encryption (K,M)->C is Symmetric encryption algorithm that takes both convergent key K and a data copy M as inputs and outputs the ciphertext C.

3) Decryption (K,C)->M is a decryption algorithm that takes the both ciphertext C and convergent key K as inputs and outputs the original data copy M.

4) Tag Generation (M)->T(M) is a Tag generation algorithm that maps the original data copy M and then outputs a tagT(M).

- A user or data owner derives a convergent key from each original data copy and encrypts the data copy with the convergent key.
- It derives a tag to the data copy such as tag will be used to detect duplicates.
- To check the duplicates, the user first sends the tag to the server side to check the identical copy has been already stored.

C. Proof of owernship

The notion of proof of ownership enables users to prove their ownership of data copies to the storage server.

D. Identification protocol

- It exits in two phrases proof and verify.
- In this stage of proof, a prover/user can demonstrate his identity to verify by performing some identification proof to his identity.

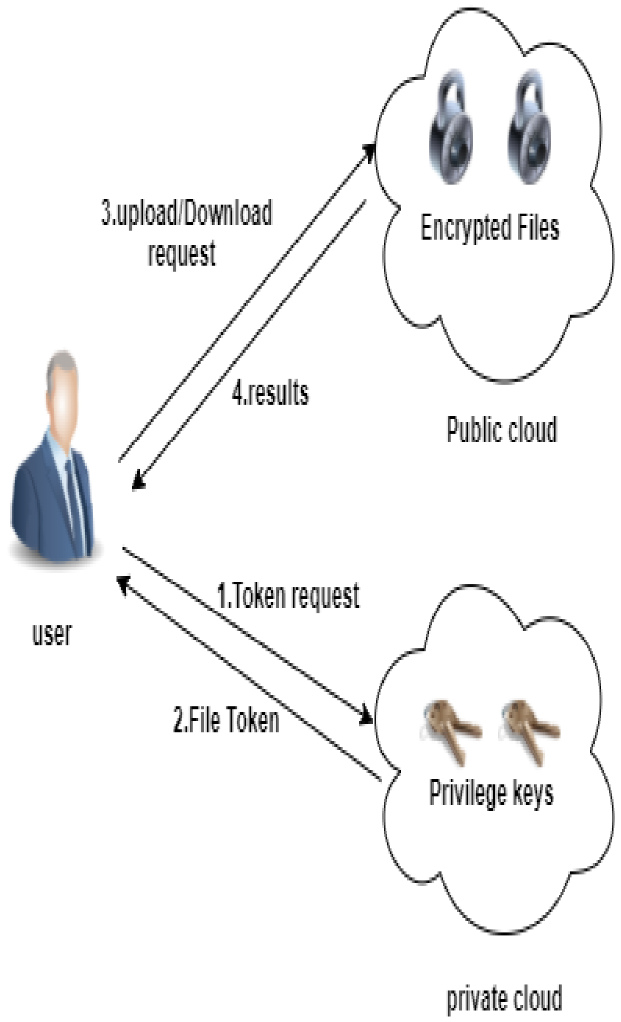

Fig.1.Architecture for Authorized Deduplication

IV. SYSTEM MODEL

A. *Hybrid Architecture for Secure Deduplication*

At a high stage, our environment of curiosity is an enterprise network, such as a group of affiliated customers (for illustration, employees of a enterprise) who will use the S-CSP and store data with deduplication technique. On this atmosphere, deduplication can be on the whole utilized in these settings for knowledge backup and disater recovery applications while commonly lowering space for storing. Such systems are wellknown and are normally more suitable to person file backup and synchronization purposes than richer storage abstractions. There are three entities outlined in our system, that is, customers, confidential cloud and S-CSP in public cloud. The S-CSP performs deduplication by checking if the contents of two records are the equal and shops simplest considered one of them.

- Each privilege is represented in the type of a short message known as token. Every file is related to some file tokens, which denote the tag with exact privileges.
- A consumer computes and checks duplicate token to the public cloud for authorized duplicate check.

Customers have access to the confidential cloud server, a semi-depended on third party so that it will support in performing deduplicable encryption with the aid of generating file tokens for the requesting customers. We will give an explanation for additional the position of the private cloud server under. Customers are additionally provisioned with per-user encryption keys and credentials (eg., person certificates). We will be able to handiest consider the file-stage deduplication for simplicity. In a different word, we refer a data copy to be a entire file and file-degree deduplication which eliminates the storage of any redundant documents. Clearly, block-level deduplication can be without problems deduced from file stage deduplication, which is similar to [12]. Primarily, to add a file, a user first performs the file-degree duplicate investigate. If the file is a duplicate, then all its blocks need to be duplicates as well; in any other case, the user additional performs the block-degree and identifies the unique blocks to be uploaded. Each data copy (i.e a file or a block) is related to a token for the duplicate check.

1) *S-CSP*

This is an entity that provides a data storage provider in public cloud. The S-CSP provides the data outsourcing carrier and stores knowledge on behalf of the users. To lessen the storage cost, the S-CSP eliminates the storage of redundant data through deduplication and maintains handiest specified information. We anticipate that S-CSP is continuously on-line and has abundant storage ability and computation power

2) *Data Users*

A user is an entity that wishes to out-source data storage to the S-CSP and access the information later. In a storage method assisting deduplication, the consumer handiest uploads precise information however does not upload any duplicate knowledge to save lots of the add bandwidth, which is also owned via the equal user or one of a kind customers. In the authorized deduplication process, each and every consumer is issued a suite of privileges in the setup of the approach. Each file is protected with the convergent encryption key and privilege keys to recognize the approved deduplication with differential privileges.

3) *Private Cloud*

Compared with the traditional deduplication architecture in cloud computing, this is a new entity introduced for facilitating user’s secure usage of cloud service. Specifically, since the computing resources at data user/owner side are restricted and the public cloud is not fully trusted in practice, private cloud is able to provide data user/owner with an execution environment and infrastructure working as an interface between user and the public cloud. The private keys for the privileges are managed by the private cloud, who answers the file token requests from the users. The interface offered by the private cloud allows user to submit files and queries to be securely stored and computed respectively.

V. PROPOSED MODELS

A. *Cloud Service Provider*

- In this module, we develop Cloud Service Provider module. This is an entity that provides a data storage service in public cloud.
- The S-CSP provides the data outsourcing service and stores data on behalf of the users.
- To reduce the storage cost, the S-CSP eliminates the storage of redundant data via deduplication and keeps only unique data.

Assume that S-CSP is always online and has abundant storage
Capacity and computation power.

B. *Data User Module*

- A user is an entity that wants to outsource data storage to the S-CSP and access the data later.
- In a storage system supporting deduplication, the user only uploads unique data but does not upload any duplicate data to save the upload bandwidth, which may be owned by the same user or different users.
- In the authorized deduplication system, each user is issued a set of privileges in the setup of the system. Each file is protected with the convergent encryption key and privilege keys to realize the authorized deduplication with differential privileges.

C. Private Cloud Module

- Compared with the traditional deduplication architecture in cloud computing, this is a new entity introduced for facilitating user's secure usage of cloud service.
- Specifically, since the computing resources at data user/owner side are restricted and the public cloud is not fully trusted in practice, private cloud is able to provide data user/owner with an execution environment and infrastructure working as an interface between user and the public cloud.
- The private keys for the privileges are managed by the private cloud, who answers the file token requests from the users. The interface offered by the private cloud allows user to submit files and queries to be securely stored and computed respectively.

D. Secure Deduplication System

- A several types of privacy we need protect, that is, unforgeability of duplicate-check token: There are two types of adversaries, that is, external adversary and internal adversary.

- The external adversary can be viewed as an internal adversary without any privilege.

- If a user has privilege p, it requires that the adversary cannot forge and output a valid duplicate token with any other privilege p′ on any file F, where p does not match p′. Furthermore, it also requires that if the adversary does not make a request of token with its own privilege from private cloud server, it cannot forge and output a valid duplicate token with p on any F that has been queried.

VI. RESULTS

The results are appeared in the web browser according to the different privilege records they are uploaded in the public cloud. The each and every individual record is stored in the public cloud, where the duplicate check is tested in the private cloud while uploading the same file again. The deduplication is done in the private cloud before it updated in the public cloud. The different testbed experiments that are appeared in comparing the same files and file Size.

VII. CONCLUSION

In the project permitted expertise deduplication was once to protect the information protection by using including differential privileges of purchasers within the duplicate verify. We additionally furnished a quantity of recent deduplication constructions helping accepted duplicate check in hybrid cloud architecture, where the replica verify tokens of documents are generated by the use of the personal cloud server with personal keys. Security analysis demonstrates that our schemes are at ease in phrases of insider and outsider assaults distinct in the proposed security model. We applied a prototype of our authorized duplicate check scheme and conduct testbed experiments on our prototype. Enchancing that our authorized duplicate check scheme incurs minimal overhead in assessment with convergent encryption and network transfer.

VIII. ACKNOWLEDGMENT

The feel glad to express my deep since of gratefulness to my project guide S.kranthi (Assistant professor), for her guideness and assistance in completing the project successfully. I thank one and all who have renderd help directly or indirectly in completion of this project successfully.

REFERENCES

[1]. OpenSSL Project. http://www.openssl.org/.
[2]. P. Anderson and L. Zhang. Fast and secure laptop backups with encrypted de-duplication, *USENIX LISA*, 2010.
[3]. M. Bellare, S. Keelveedhi, and T. Ristenpart. Dupless: Server-aided encryption for deduplicated storage, 2014.
[4]. M. Bellare, S. Keelveedhi, and T. Ristenpart. Message-locked encryption and secure deduplication, 2015.
[5]. M. Bellare, C. Namprempre, and G. Neven. Security proofs for identity-based identification and signature schemes. 2015.
[6]. M. Bellare and A. Palacio. Gq and schnorr identification schemes: Proofs of security against impersonation under active and concur-rent attacks. 2014.
[7]. S. Bugiel, S. Nurnberger, A. Sadeghi, and T. Schneider. Twin clouds: An architecture for secure cloud computing, 2014.
[8]. J. R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M. Theimer. Reclaiming space from duplicate files in a serverless distributed file system. 2015.
[9]. D. Ferraiolo and R. Kuhn. Role-based access controls, 2014.
[10]. GNU Libmicrohttpd. http://www.gnu.org/software/libmicrohttpd/.
[11]. S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg. Proofs of ownership in remote storage systems. In Y. Chen, G. Danezis, and V. Shmatikov, 2014.
[12]. J. Li, X. Chen, M. Li, J. Li, P. Lee, and W. Lou. Secure deduplication with efficient and reliable convergent key management, 2014.
[13]. libcurl. http://curl.haxx.se/libcurl/.
[14]. C. Ng and P. Lee. Revdedup: A reverse deduplication storage system optimized for reads to latest backups, 2014.
[15]. W. K. Ng, Y. Wen, and H. Zhu. Private data deduplication protocols in cloud storage. In S. Ossowski and P. Lecca, 2014.

[16]. R. D. Pietro and A. Sorniotti. Boosting efficiency and security in proof of ownership for deduplication. In H. Y. Youm and Y. Won, editors, 2014.

[17]. S. Quinlan and S. Dorward. Venti: A new approach to archival storage. 2014.

[18]. A. Rahumed, H. C. H. Chen, Y. Tang, P. P. C. Lee, and J. C. S. Lui. A secure cloud backup system with assured deletion and version control, 2014.

[19]. R. S. Sandhu, H. L. Feinstein, and C. E. Youman. Role based access control models, IEEE computer 1996.

[20]. J. Stank, A. Somottii, E. Androula, and L. Kencl. A secure data scheme deduplication cloud storage, 2014.

[21]. M. W. Storer, K. Greenan, D. D. E. Long, and E. L. Miller. Security data deduplication, 2014.

[22]. Z. Wilcox O'Hearn and B. Warner. Tahoe: The latest authority security, 2015.

[23]. J. Xu, E.-C. Chang, and J. Zhou. Weak leakage-resilient client-side deduplication of encrypted data in cloud storage,2014.

[24]. J. Yuan and S. Yu. Secure and constant cost public cloud storage auditing with deduplication, 2015.

.