International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 9, September 2016)

A Implementation of Hybrid Cloud Approach with Elimination of Equivalent File Data

Danej Pooja[1], Gorade Varsha[2], Gaikwad Shital[3], Salmuthe Priyanka[4], Prof. V.R. Sonawane[5]
Computer Engineering, SND College Of Engineering & Research Center, Yeola, India

***Abstract*— Now a days the cloud computing is very popular technology used for the data processing. The cloud computing is deals with the data stored on a cloud. It provide a efficient and secure data storage facility to normal computer user. Again there is major concern about the security , for this the convergent key mechanism is used, there is also chances of the leakage of the data because there is a very high data traffic on cloud, so in propose system it use the encryption and decryption technique for providing the data security and confidentiality, it provide the unique key for each user. But on the cloud the data redundancy is occur because that is consume extra bandwidth and more memory. So in propose system we eliminate that redundancy (about data). And also used encryption and decryption technique, therefore data confidentiality is mentioned.**

***Keywords*— Hybrid cloud, privilege mechanism, convergent key, duplicate check, security.**

I. INTRODUCTION

Now days cloud computing is very popular technology used for dealing with the day to day life. The cloud is a data where the data is stored on a one large database and multiple users can access that data anywhere, anytime. There are several types of cloud data storage. In propose system we use the Hybrid cloud approach. In Hybrid cloud there is a combination of two clouds, public cloud and private cloud. Using Hybrid cloud approach data will stores in effective manner and it improves the performance of system[4].

As we know cloud computing is a recently used and vast data processing technology, therefore there is a chance of duplication of data on a cloud, it consumes extra memory and hence it degrades the performance of the system. So instead of keeping multiple data copies of same data on a single cloud, we eliminate those duplicate data by keeping single copy(deduplication). decryption techniques for provide the data security and confidentiality. On the cloud multiple users upload and download data to and from the cloud simultaneously. Therefore it needs to maintain a unique identity between the different users. Also there may be a chances of data leakage, we know there is a very high data traffic on cloud computing for avoiding the same problem we provide the key(token) privilege mechanism[5,3].

II. SURVEY OF LITERATURE

By referring various papers for knowing the detailed idea about proposed system .also trying to understand multiple system related to proposed system.

Secure storage-Now days era of information get increased in corporation sector there is a requirement no. of resources daily basis and which is not use further after used, so to buy that resources temporary is possible using cloud computing. It is secure storage system. For example mozy, drop box, torna informatics, amazons[7].

Secure Authentication-Prevent the data against the unauthorized person, improves the security of system for archive such security the privileged mechanism were used.

Convergent Key encryption- The convergent key encryption for insider and outsider attacker .for monitoring such attacks' and preventing data against them [1, 3, 5].

By the survey of previous system we observe **some** problems

In previous system deduplication cannot prevent the privilege private key sharing among the users. User will issue some private key for accessing the same file so collude occurred.

In previous system it does not provide PoW or user identification. One of the critical designs is that the traditional conversion encryption system can only protect semantic security of unpredictable file.

To solved the problem of previous system we proposed the novel system which supporting authorized duplicate check in it .we introduced hybrid cloud architecture. The private key for privilege will not issue to the user directly which will be kept and manage by the private cloud server instead.

The private cloud server also check user identity before issuing the corresponding file token to the user and authorized duplicate check perform on public cloud before file on it[1,3,5].

III. PROPOSED SYSTEM

Here we design and implement a new system "Eliminate equivalent data using cloud approach". The major goal of this system is to avoid duplication of replicated data (repeated data) for that we use convergent key mechanism.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 9, September 2016)

Convergent key

Convergent key encryption technique allows the data confidentiality in deduplication. A user (data owner) derives a convergent key from each and every original data copy and encrypts the data copy with the convergent key. In addition, a user also derives a tag for data copy such that tag will be use to detect duplication. Here, we assume that the tag correctness property holds, if two data copies are same, then their tags are the same. To detect duplication, the user first sends the request to the server side to check if the identical copy has been already stored, note that both the convergent key and the tags are independently derive , and the tag cannot used to deduce convergent key and compromise data confidentiality. Both the encrypted data copies and its corresponding tags will be stored on a server side. Formally a convergent encryption scheme can be defined with a four primitive function.[5]

KetGence(M) = K is key generation algorithm that maps a data copy M to a convergent key K.

EnCce(K,M) = C is the symmetric encryption algorithm that take both the convergent key K and data copy M as a input and outputs a cipher text C.

Decce(K,C) = M is the decryption algorithm that take both the cipher text C and convergent key K as a input and then output the original data copy M. **TagGen(M)** = T(M) is the tag generation algorithm that maps the original data copy M and outputs a tag T(M).

Table 1.

Acronym	**Description**
S-CSP	Storage-Cloud Service Provider
PoW	Proof of Ownership
(pkU,skU)	Users public key & secrete key pair
pU	Privilege set of user U
kF	Convergent encryption key for file F
pF	Specified privilege set of file F
ϕ'F,p	Token of file F with privilege P

Symmetric Encryption

Symmetric encryption uses a common secret key κ to encrypt and decrypt information. A symmetric encryption scheme consists of 3 primitive functions[5]:

• KeyGenSE(1λ) → κ is the key generation algorithm that generates κ using secure parameter 1λ.

• EncSE(κ,M) → C is the symmetric encryption algorithm that takes the secret κ and message M and then outputs the cipher text C.

• DecSE(κ,C) → M is the symmetric decryption algorithm that takes the secret κ and cipher text C and then results will be the original message M.

Proof of Ownership (Identification Protocol)

The idea of proof of ownership (PoW) enables users to prove their ownership of data copies to the storage to the server. Specifically, proof of ownership is implemented like interactive algorithm run by a user and a server. The server derives a short value ϕ(M) from a data copies M. To prove the ownership of the data copy M, the user needs to send ϕ′ to the server like ϕ′ = ϕ(M). The formal security dentition for proof of ownership raw follows the threat model in a content distribution network, where an attacker does not know the whole files, but has accomplices who have the file. The accomplices follow the "bounded retrieval model", such that they helps to the opponent obtain the files, subject to the constraint that they must send little bits than the starting min entropy of the files to the attacker[5,3].

An identification protocol Π is described in two phases: Proof and Verify. In first stage of Proof, a user U can convey his identity to a server by performing some identification proof related to his identity. The input of the user is his private key skU that is sensitive information like private key of a public key in his certificate or credit card number, etc. that he would not like to share with the other users. The server performs the verification with input of public information pkU related to skU. At the end of the protocol, the server outputs either accept or not to denote whether the proof is passed or fail.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 9, September 2016)

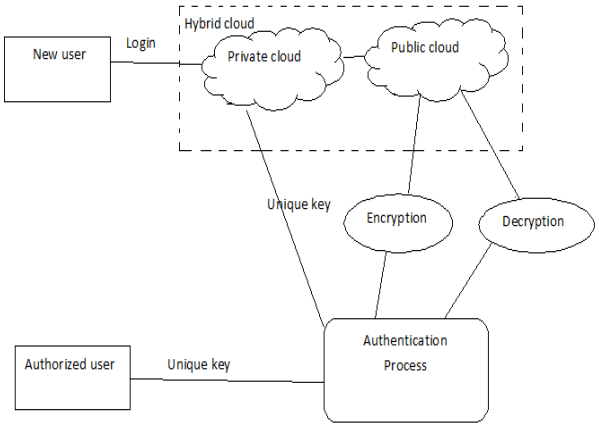

Fig.1.: Architecture of Secure Authentication System

In propose system architecture main purpose is to avoid the duplicate data which is stored on a public cloud. This system also provide secure authorized user by using some strategy. Such as symmetric key encryption and convergent key mechanism. In above system architecture we are using two cloud (hybrid cloud) i.e. public and private cloud. In private cloud side user registration process is carried out in that the unique key is provided to the user (File token). In other side public cloud act as a storage cloud service (provider) to symmetric key technique are used such as symmetric encryption and decryption. from the public cloud we can upload file and also we can download file on it with its uniqueness of file(data).

• S-CSP. It is an entity that provides a storage service in public cloud. The S-CSP provides the data outsourcing service and stores data on profit of the users. To reduce the storage cost, the S-CSP remove the storage of repeated data via deduplication and maintain only unique data. In propose system we assume that S-CSP is always online and has copious storage capacity and computation power.

• Data Owner. A owner is an entity that wants to out-source data storage to the S-CSP and access the data in future. One can only store the unique data ,the duplicate data cannot be upload its result extra bandwidth is save, which may be owned by the same user or different users. In the authorized equivalent data identification system, each user is to give a set of privileges in the setup of the system. Each file is secured with the convergent encryption key and privilege keys to realize the authorized equivalent with differential privileges. hardware- based security features to implement a remote execution environment trusted by the users.

• Private Cloud. Compared with the previous equivalent identification data architecture in cloud computing, this is a new entity introduced for facilitating users secure usage of cloud service. Specifically, since the computing resources at data owner side are secured and the public cloud is not fully trusted in practice, private cloud is able to provide data owner with an execution environment and infrastructure working as an interface between user and the public cloud. The unique keys for the privileges are control by the private cloud, who answers the file token requests from the users. The interface offered by the private cloud allows user to resign files and queries to be securely stored.

Notice that, this is a new architecture for data eliminating redundancy in cloud computing, which includes two clouds. Actually, this hybrid cloud setting has attracted more and more attention now days. For example, an enterprise might use a public cloud service, such as drop box, for to hide data, but continue to maintain in-house storage for operational customer data. Alternatively, the secure private cloud could be a cluster of virtualized crypto graphic coprocessors, which are offered as a service by a TP (Third party) and provide the necessary.

IV. Algorithmic Strategy

In this paper we propose a new equivalent file identification system support the following:

• *Differential Authorization:* Each and every authorized user is able to get his/her separate token of his files to perform duplicate check based on his privileges. In this assumption, any user cannot generate a token for duplicate check out of his privileges or without the grant from the private cloud server.

• *Authorized Duplicate Check:* Authorized user is able to use his/her individual unique keys to generate query for certain file and the privileges he/she owned with the assist of private cloud, while the public cloud performs duplicate check instantly and tells the user if there is any duplicate. It considered security requirement , in this paper lie in two folds, including the security of files token and data files. For the security of files token, two things are define as unforgeability and in-distinguishability of files token. The details are given below.

• *Duplicate-check token:* Unauthorized users without appropriate privileges or file should be prevented from getting the file tokens for duplicate check of any files stored at the S-CSP. The users are not granted to collude with the public cloud server to break the unforgeability of files tokens. The duplicate check token of a user should be issued from the private cloud in this system.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 9, September 2016)

• *In-distinguishability of files token:* It requires that any user without querying the server of private cloud for some file token, he cannot get any helpful information from the token, which contains the files information or the privilege information.

• *Data Confidentiality:* Unauthorized users without a appropriate privileges or files, including the S-CSP and the private cloud should be prevented from access to the underlying plaintext stored at S-CSP. In other word, the goal of the adversary is to retrieves and recovers the files that do not belong to cloud server. In our system, compared to the previous definition of data confidentiality based on convergent encryption, a higher level confidentiality is define and achieved.

Types of Algorithms:

In this scheme, we use two types of algorithms,

1) For file upload.
2) For file download.

For Uploading A File

BEGIN

Step1- Read file

Step2- Cloud server verifies for duplication

Step3- Sends duplication result whether the file already exists or not

Step4- If the file not exist

4.1 Display message "file does not exist"

Step5- Then uploads the file

Step6- If the file is already exist

6.1 then Display the message "file already exist"

END

For Downloading A File

BEGIN

Step1- Read the file

Step2- Cloud server checks for duplication

Step3- Sends duplication response whether the file already exists or not

Step4- If the file exist

4.1 Display "file exist"

Step5- then it downloads the file

Step6- If a file does not exist

6.1 Display message "file does not exist"

END

Our implementation of the Client provides the following function calls to support token generation and de-duplication along the file upload process.

• FileTag(File) - It computes the value of SHA-1 hash as File Tag from file.

• TokenReq(Tag, UserID) - It requests the Private Cloud for File Token generation with the File Tag and User ID.

• DupCheckReq(Token) - It requests the Storage Server for Duplicate Check of the File by sending the file token received from private cloud server.

• ShareTokenReq(Tag, {Priv.}) –It requests to the Private cloud Server to generate the Share File Token with the File Tag and Target Sharing Privilege Set.

• FileEncrypt(File) - It encrypts the File with Convergent Encryption using the 256-bit AES algorithm in cipher block chaining mode, where the convergent key is from SHA-256-Hashing of the file.

• FileUploadReq(FileID, File, Token) - if the file is Unique, it uploads the File Data to the Storage Server and updates the File Token stored.

• DupCheck(Token) - It searches the File to Token Map for Duplicate.

• FileStore(FileID, File, Token) - It stores the File on cloud and updates the Mapping.

Our implementation of the Private Server includes corresponding request handlers for the token generation and maintains a key storage with Hash Map.

• *TokenGen(Tag, UserID)* - It loads the related privilege keys of the user and generate the token with the HMAC-SHA-1 algorithm.

• *ShareTokenGen(Tag, {Priv.})* - It generates the share token with the coherent privilege keys of the sharing privilege set with HMAC-SHA-1 algorithm.

• *DupCheck(Token)* - It searches the File Token Map for Duplicate.

• *FileStore(FileID, File, Token)* - It stores the File on cloud and updates the mapping.

V. Conclusion

We present this system in which we provide a security and deduplication of replicated data over cloud.

The main idea of this system is to secure the users data using convergent keys mechanism and also avoid the redundant data which is stored on the cloud and for preventing such problem we used symmetric encryption-decryption and convergent key mechanism.

REFERENCES

[1] Agrawal1, Nikhil O. "Secure Deduplication and Data Security with Efficient and Reliable Convergent Key." IJARCCE (2015): 4.

[2] Aparna Ajit Patil, Asst. Prof. Dhanashree Kulkarni. "Block Level Data Duplication on Hybrid Cloud Storage System." IJARCSSE (2015): 6.

[3] Bhushan Choudhary, Amit Dravid. "A Study On Authorized Deduplication Techniques in Cloude Computhing." IJARCET (2014): 4.

[4] JADAPALLI NANDINI, RAMIREDDY NAVATEJA REDDY. "IMPLEMENTATION OF HYBRID CLOUD APPROACH FOR SECURE." IRJET (2015): 10.

[5] Jin Li, Yan Kit Li, Xiaofeng Chean, Patrick P.C. Lee, Wenjing Lou. "A Hybrid Cloud Approach For Secure Authorized Duplication." IEEE (2014): 12.

[6] M. Bellare, S. Keelveedhi, and T. Ristenpart. "Message-locked encryptionandsecurededuplication." IJES (2013): 6.

[7] Mihir Bellaie, Shriram Keelveedhi, Thomas Ristenpart. "DupLESS:Server-Aided encreption For Duplicate Storage." IJARCET (2013): 16.

[8] Ms. Madhuri A. Kavade, , Prof. A.C.Lomte. "A Literature Survey On Secure De-Duplication Using Convergent Encryption Key Management." IJECS (2014): 4.

[9] Pooja S Dodamani, Pradeep Nazareth. "A Survey on Hybrid Cloud with De-Duplication." IJIRCCE (2014): 10.

[10] Prajakta Patil, Mr. Anilkumar Warad. "A Survey on Data De duplication Techniques ." IJCTA (14): 4.

[11] V.R. Sonawane, D.R. Rao. "An Efficient Hybrid Clustering Approach for Multi Version XML Document ." JTAIT (2015): 6.

[12] "An Optimicity Approach for Clustering multi version XML Docuoments using Compressed Delta." IJECE (2015): 8.

[13] V.R.Sonawane, D.R.Rao. "XML Document Management Tools Survey." EIJ (2015): 4.