ijecs open access www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 3 Issue 5 May, 2014 Page No. 5526-5632

Security and Privacy in Private Cloud Storage

V.B. Maral[1], Mehul Sanghvi[2], Pankaj Rana[3], Wajeda Pathan[4], Heena Parihar[5]
Professor, Dept. of Computer, K J College of Engineering and management Research Pune, India[1]
Dept. of Computer K.J College Of Engineering Pune [2,3,4]

***Abstract*—Cloud Storage is a new emerging technology in which users can outsource their data on cloud and access it remotely from anywhere, thus the user can be free from the burden of data storage and its maintenance. But it is true that the data outsourced on cloud storage is not under the possession of user, so there is always a risk of integrity of data. The Cloud Storage should be such that the user should store the data on cloud as they are storing data on local system without worrying about data integrity. So for this purpose an external third party auditor (TPA) is introduced which performs the task of public auditing and users can rely on TPA for security of outsourced data. TPA performs the auditing of data, but this should not bring new risk to user's data security by checking the contents of data while performing auditing. In this paper, we propose a scheme privacy preserving public auditing which stores the user's data securely on the cloud. The proposed scheme is highly efficient and provably secured against both cloud service providers and external third party auditor (TPA).**

Index Terms—Privacy preserving, public auditability, cloud computing.

I. INTRODUCTION

Cloud Computing has been seen as the next-generation trend of IT industries, because of its various advantages: omnipresent network access, on-demand self-service, location independent resource pooling and usage-based pricing. With the help of cloud computing we can use many cheaper and much powerful set of processors, together with the software as a service also called as SaaS computing architecture, which are changing data centers into pools of computing service. Cloud in cloud computing is nothing but a set of software, hardware, storage, networks and services which when combined delivers various aspects of computing as a service to the users. Various resources, software and also information are provided to users on demand.

The users are charged according to the resources they use, thus making them free from maintenance of these resources. Using cloud computing clients can access various IT resources with the help of which they can deploy new services, resources or applications. The main concept of cloud computing is to free the users from the burden of processing by providing an interface which facilitates an efficient processing mechanism. All this can be simply done through a simple internet connection using any standard browser.

As cloud service providers (CSP) are different administrative party, user loses his all control over the data after outsourcing the data. The user's data which is outsourced on cloud is not safe i.e. there may be data integrity problems from the cloud service providers in case cloud service providers might behave unfaithfully for their own benefits or profit. For example a cloud might delete a data of user which is rarely used just to gain more space and also there are chances of some data getting corrupted. Some hackers may also attack the data stored in cloud and to protect their reputation they may not tell the users about the data later or being hacked. In such cases users cannot totally depend on cloud storage for security or integrity of their data. It is impossible to detect any alteration done to the data as the user may outsource bunch of data to the cloud. So it becomes necessary for user to detect any alteration done to the data and for verifying the data the user has to download the copy of data every time. This is wastage the user's time and also the bandwidth of the network.

So to remove the burden of all the auditing process by user itself, user may rely on Third Party Auditor (TPA) for data integrity who are professional in auditing the data for any possible alteration on demand of users on their data. The word public auditing will audit the data publicly i.e different users.Therefore, the question rises as to how we can achieve the privacy-preserving third party auditing protocol, is the problem we are going to tackle in this paper. The user may Encrypts the data and then store outsources data into the cloud, but this does not ensures the security to data and also does not ensures the privacy to your data

II. RELATED CONCEPTS ABOUT CLOUD

A. DEPLOYMENT CLOUD MODELS

- Public cloud: In this model the cloud infrastructure that is created is made available to a large industry group and also the general public people and it is provided by single service provider selling its cloud services.
- Private cloud: In this model the cloud infrastructure is operated solely and only for an

organization or by a group. One of the main advantages of this model is its compliance, security and QoS.

- Community cloud: In this model the cloud infrastructure may be shared by several organizations or by several groups and it supports only a specific community that has shared concerns like policy, security requirements.
- Hybrid cloud: The cloud infrastructure is a combination of two or more types of clouds. It also enables data application portability through load balancing between these clouds.

B. CLOUD CHARACTERISTICS

- On demand service: Cloud is service pool and large resource with which you can get service or resource whenever or wherever you need by paying the amount which you have utilised.
- Ubiquitous network access: Cloud service provides cloud services that are present everywhere though standard terminal like laptops, mobile phones and through personal digital assistants.
- Easy use: Most cloud service provider's offers internet based interfaces which are nothing but application program interfaces also called as api so that the user or client can easily use the provided cloud services by the cloud service providers.
- Business model: Cloud is a business model because it is pay per use i.e. the user is only required to pay for the service of resources which he uses.
- Location independent resource polling: The cloud service provider's computing resources are deployed to serve multiple users with which the user can access the different virtual and physical resources dynamically from any location.

C. CLOUD SOLUTIONS

- Infrastructure as a service: It delivers a platform virtualization environment to the user as a service rather than purchasing software, servers or data canters.
- Software as a service: It is software that is deployed to run behind a firewall in LAN or PC or it is deployed over internet.
- Platform as a service: This is a kind of cloud computing service that provides development environment as a service. This can be used as the middleman's equipment with which one can develop their own program and deliver that program to the users through internet.

III. SOFTWARE AS A SERVICE (SAAS)

Software as a Service model provides the capability to user to to use the service provider's applications via internet or LAN through user's PC or mobile. The applications are accessible from various devices of user through a thin interface such as a web browser or a web application. The user does not need to worry for managing or controlling the underlying cloud infrastructure which may include servers, network, storage, operating systems.

In our system we are using Software as a Service model which would provide a storage service to the user through an application which would store the user's data.

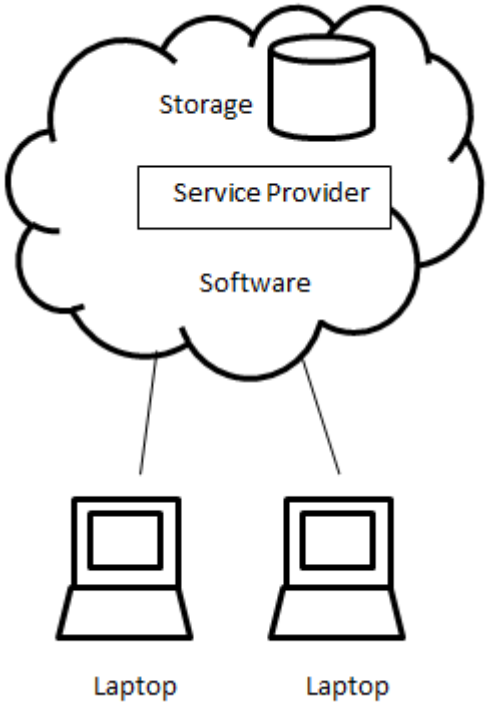


Fig: SaaS Architecture

There are a number of advantages of SaaS which is beneficial to organisations and personal users like:

- No additional hardware costs: The Required processing power to run the applications is provided by the cloud service provider.
- No initial setup costs: The applications are ready to use as soon as the user pays or subscribes it.
- Pay for what you use: If a piece of software is only needed only for a small period then it is paid for over that period and it's subscriptions can usually be halted or suspended at any time after there is no use of it.
- Usage is scalable: If a user decides they need more services or any additional storage then they can access those services on demand without needing to install any new software or hardware.
- Updates are automated: Whenever there is an update available, it is provided to existing customers, more often free of charge. No new software installation will be required as it already deployed by the cloud service provider.
- Cross device compatibility: SaaS applications can be accessed through any internet enabled device (internet enabled PC, Mobile and Tablet) which makes it ideal and convenient for those who use a number of devices and for those user who don't always use the same computer.
- Accessible from any geographical location: A deployed application can be accessed from any place

with any internet enabled device rather than being restricted to particular installations on individual computers.

IV. Existing System

The existing cloud storage service involves two different entities-

- The cloud user (U) is an entity that has large amount of files which are to be outsourced i.e. to be stored on cloud.
- The cloud server (CS) is managed and maintained by the cloud service provider (CSP).

The user outsources his data on the cloud and there are chances of data integrity threats, the threats may arrive from internal sources or external sources like- hackers, management errors or even bugs. The existing system does not ensure data integrity and so the users can't rely on cloud service provider for their important outsourced data.

V. Goals

For ensuring integrity of the data outsourced, the system should achieve certain parameters:

1. Public auditability: In this type of auditing the TPA chooses random users and their random files and audits them, the TPA audits the user's data without retrieving the copy of data. The TPA checks the correctness of the data outsourced.
2. Privacy Preserving: The system should take care that the external third party auditor who is auditing the data, should read or modify the data which is outsourced on the cloud by the user.
3. Correctness: The data outsourced on the cloud is neither modified nor read by cloud service provider.
4. Dynamics: The user can dynamically upload, download, modify or delete the data.
5. Lightweight: The auditing process must be done by the TPA in such a way that there should be no minimum computation and communication overhead.

VI. Proposed System

The proposed cloud storage service involves three different entities-

- The cloud user (U) is an entity that has large amount of files which are to be outsourced i.e. to be stored on cloud.
- The cloud server (CS) is managed and maintained by the cloud service provider (CSP).
- The external third party auditor (TPA) who audits the data.

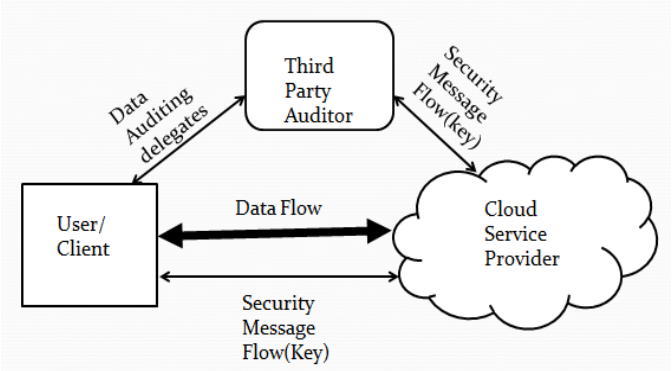


Fig. 1

In the proposed system an entity called TPA is added, who performs the task of public auditing. The TPA chooses random file of the users' and audits them. After the process of auditing is done the TPA notifies the user whether his outsourced data is modified or not. The correctness of data is ensured by encrypting the data before storing it onto the cloud. For encryption, first of all the data is partitioned and then each block is encrypted. Using this cloud service provider and the TPA cannot read the contents of the data. All the operations – upload, download, delete, modify can be performed dynamically.–--

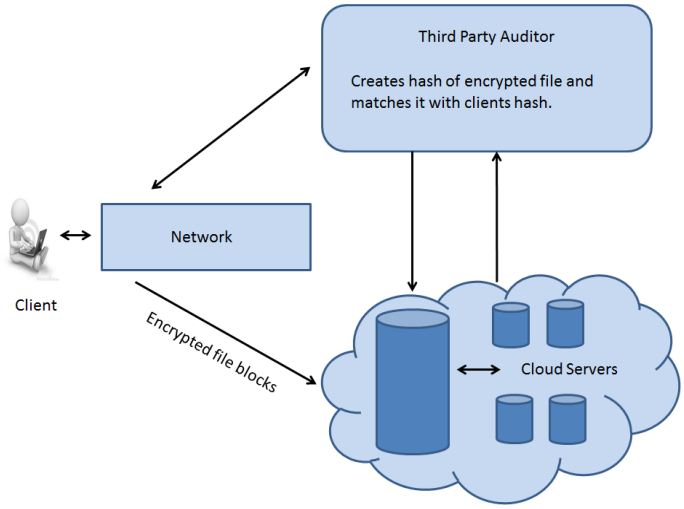
The system ensures that the local copy of data is not exposed to the TPA; only the metadata required for auditing is retrieved, thus attaining privacy preserving.

VII. System Schema

A. Privacy Preserving Public Auditing Scheme

Our scheme uses Advanced Encryption Standard (AES) for encryption and decryption of the data to be outsourced and before outsourcing the data on cloud storage. The client partitions the file into small blocks i.e. small partitions of data and then encrypts these partitions. After these partitions are encrypted, a digital signature or hash for each partitions is computed, which is required for verifying the correctness of data. This Hash or digital signature is computed using Secure Hashing Algorithm 1 (SHA1).

The User stores the encrypted partitions on cloud storage and the hash of each partition on the TPA database which is later used for verifying the integrity of the data outsourced. Later TPA challenges the CSP for verifying the data demanding random encrypted partitions of random files of random user. Then TPA now computes the hash of these encrypted partitions and compares it with the previously obtained hash from the user. If the hashes are similar then TPA notifies the user that the outsourced partition is safe and if the hashes do not match a message that the partitions have been altered is sent to the user.



The TPA cannot view the contents of partitions as these partitions are encrypted by the user on user's machine and only the user has the key to decrypt the partition. This thus implements our scheme Privacy Preserving Public Auditing Scheme, where Public Auditing is implemented using SHA1 and Privacy Preserving is implemented using AES.

Mathematical Model

Notation and Preliminaries

Let S= {U, F, Db, Sc, TPA, P}

1. Where U is set of users $U_1, U_2, \ldots U_n \in U$
2. F is a set of files $f_1, f_2, \ldots f_n \in F$
3. Db is a set of data tables used to store user information, file information and partition details.
4. Sc is data storage server, where the files are stored.
5. TPA is auditor who audits the data.
6. P is a set of file Partition where $P_1, P_2, \ldots P_n \in P$

1. Client Side

U=RegisterUser(uid,pwd,name,addr,country,contact,email)
Auth[Y:N]=login(uid,pwd')
pwd'=MD5(pwd)
P=partitionFile(file F, partition size)
EP=encryptPartition(P,key)
H=Hashfiles(P)
Upload(EP,Sc)
Upload(H,TPA)

2. TPA Side

R=RandomPartition(Sc)
C=GenHash(R)
O=ObtainHash(EP,TPA)
L[Y:N]=verifyHash(C,O)

3. Server Side

RespondTPA(EP)

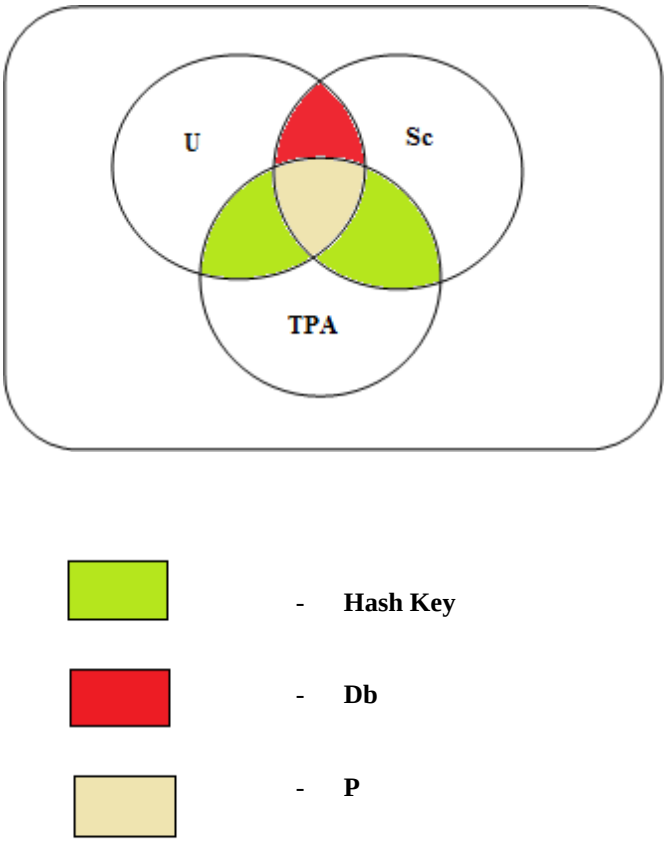


Fig. Venn diagram

VIII. IMPLEMENTATION DETAILS

In the proposed privacy preserving public auditing scheme, we have done following things:

1. File Partition
2. Encryption
3. Hash Calculation
4. Decryption

1. File Partition

The user file cannot be uploaded as whole, so we have to upload the file in small blocks. The block size can be user defined i.e. 1024 bytes, 2048 bytes, 4096 or any other size the user wants. We have decided to partition the file in blocks of size of 4096 bytes, so that the transfer of file blocks over the network becomes fast.

2. Encryption and Decryption

The encryption and decryption of file blocks is done using advanced encryption standard (AES) algorithm, it is by far the most widely used algorithm for encryption and decryption. It is computationally more fast and safer than DES. AES changes the contents of file into cipher text, so that the contents are unreadable.

AES Algorithm:

AES is the Advanced Encryption Standard, developed by United States government standard algorithm for encrypting data.

AES is block cipher each having block size of 128 bits.

AES KEY TABLE

Sr No.	Key Length	Keyname	Rounds
1	128	AES-128	10
2	192	AES-192	12
3	256	AES-256	14

AES performs encryption by executing following functions:
•SubBytes()
•ShiftRows()
•MixColumns()
•AddRoundKey()
SubBytes(), ShiftRows(), MixColumns() functions are used to provide security from cryptanalysis by creating dispersed pattern of plaintext in cipertext and by creating a relationship between the plaintext and ciphertext. AddRoundKey() function is maily used to encrypt the data..

SubBytes():SubBytes() is a function that adds confusion to the data by processing each byte through a predefined S-Box.

ShiftRows(): ShiftRows() is a function that creates diffusion by mixing data within rows. The 0^{th} Row of the State is not shifted, 1^{st} row is shifted by 1 byte, 2^{nd} row is shifted by 2 bytes, and 3^{rd} by 3 bytes.

MixColumns(): MixColumns() function also creates diffusion by mixing required data within columns. The 4 bytes of each column in the State are treated as a 4-byte number and it is transformed to another 4-byte number through finite field mathematics.

AddRoundKey(): Main encryption on the data is performed in the AddRoundKey() function, where each is XORed with the subkey in the State. Using key expansion schedule the subkey is derived from the key.

This is only one round of encryption. Similarly we have to perform 10 rounds for 128 bit key length, 12 rounds for 196 bits.

AES Decryption: Decryption occurs through the function AddRoundKey() , plus the inverse AES functions invShiftRows(), InvSubBytes(), and InvMixColumns()

3. Hash Calculation

The scheme requires hashing two times, once when the user password has to be converted and second when the digital signature of each file blocks has to be calculated.

User when enters the password it cannot be exposed to CSP, using message digest 5 (MD5) algorithm the text password is converted into hash and then stored in CSP database and later for authentication, the hash values are matched.

For digital signature calculation of each file blocks secure hashing algorithm 3 (SHA-3) is used. Each encrypted file block is taken and its digital signature is calculated and then these values are stored in TPA. Later TPA also calculates hashes of file blocks to verify the status of file blocks using SHA algorithm.

MD5 Algorithm

The **MD5** message-digest algorithm is a cryptographic hash function which produces a 128-bit i.e. 16-byte hash value. MD5 has been used in a wide variety of cryptographic applications like password authentication and is also commonly used to verify data.

Step 1: Append padded bits:The message is padded so that its length is congruent to 448, moldulo 512.

Step 2: Append length:A 64-bit representation of b bit message is appended to the result of the previoud step.The resulting message has length of multiple of 512 bits.

Step 3: Initialize MD buffer:A four-word buffer (A,B,C,D) is used to compute the message digest each of 32 bit register.These registers are initialized to following values in hexadecimal:
Word A: 01 23 45 67
Worrd B: 89 ab cd ef
Word C: fe dc ba 98
Word D: 76 54 32 10

Step 4: Proces message in 16-word blocks:Four auxiliary functions take three 32 bit words as input and produces one 32-bit word as output.
F(X,Y,Z)=XYvnot(X)Z
G(X,Y,Z)=XZvYnot(Z)
H(X,Y,Z)=XxorYxorZ
I(X,Y,Z)=Yxor(Xvnot(Z))

Step 5: Output:The message digest produced as output is A,B,C,D.

Secure Hashing Algorithm:

SHA is a cryptographic hash function developed by the United States National Security Agency. SHA-3 produces a 160-bit hash value. A SHA-2 typically generates a hexadecimal number, 40 digits long. SHA-2 is widely used of all the existing SHA hash functions available. SHA is also known as message digest. A main reason for the publication of the Secure Hash Algorithm was the Digital Signature Standard, in which it is included. A hash cannot be decrypted back to the original text as it is not encrypted. This makes it appropriate when it is suitable to compare the hashed versions of texts, as it cannot decrypt the text to obtain the original version. In Digital signatures by encrypting the hash of a document with your private key one can sign the hash of a document and thus producing a digital signature for the required document. Anyone else can then check whether you have authenticated the text by decryption of the signature with your public key to get back the original hash again, and by matching it with their hash of the text.

Step 1: Padding: To ensure that the message has length multiple of 512 bits:
•first, a bit 1 is appended,
•next,k bits 0 are appended, with k being the smallest positive integer such that $l+1+k \equiv 448 \bmod 512$, where
L is the length in bits of the initial message,
•finally, the length $l < 264$ of the initial message is represented with exactly 64 bits, and these bits
are added at the end of the message.
The message shall always be padded, even if the initial length is already a multiple of 512.

Step 2: Block decomposition:
For each block $M \in \{0,1\}^{512}$, 64 words of 32 bits each are constructed as follows:
•the first 16 are obtained by splitting M in 32-bit blocks
$M = W_1 k W_2 k \cdots k W_{15} k W_{16}$
•the remaining 48 are obtained with the formula:
$W_i = \sigma_1(W_{i-2}) + W_{i-7} + \sigma_0(W_{i-15}) + W_{i-16}, 17 \le i \le 64$

IX. TESTING AND EVALUATION

Security Analysis:

Security Analysis is done on the system to evaluate the security of the proposed process by analyzing it.

In Storage Correctness Guarantee we need to check that whether the cloud server cannot generate valid response toward TPA without faithfully storing the data.

In Privacy Preserving Guarantee makes sure that TPA cannot derive user's data and contents from the file or the information collected during auditing process. The back patching technique is used for the random oracle model.

Performance Analysis:

Performance analysis gives the knowledge of the systems result. The auditing mechanism is done between the TPA and the some cloud service storage were the users data is to be outsourced. Efficient auditing is also taken into account in the performance analysis of the system.

Cost of Privacy-preserving Guarantee: This performance analysis gives the estimation of the cost in terms of basic Cryptographic operations, auditing phase computation and the server side computation. Also communicational overhead is taken into consideration for the proposed system.

Testing Phase:

Types of testing carried out are given below:

Unit testing

A. Unit testing test the design for validation of program that it functions properly so that it produces desired output for the system .Unit testing is done on each module of the system. Unit testing is applied on users, the cloud service provider and the TPA. By encrypted partition files in user module , the CSP gets the encrypted file which are partitioned and digital signature in TPA unit is generated. Tools used for unit testing are JUnit – Java unit testing framework,Cactus and Struct Test Case.

Integration testing

Integration Testing is designed to test integrated software elements to know if the system works and runs correctly. Integration testing is event driven and is related to the outcome of the program on the screens. Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.the tools used for integration testing is JsTestDriver.

Output: All the test cases mentioned above passed successfully. No defects encountered in the system.

Functional test

Functional testing is done for checking the proper functionality of system and test the technical aspects carried out in the process of the system. Tool used for functional testing is Jetif.

Functional testing is carried out on various functions of system as given below:

Input:

1. Valid Input: Identified classes of valid input are accepted. Authentication for valid users is provided.
2. Invalid Input: Identified classes of invalid input are rejected. Unauthorized users are rejected.

Functions:

Identified functions are to be executed. File transaction process is tested by the TPA for integrity of data and authentication of client. System procedures and interfacing systems or procedures must be invoked.

Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements. Acceptance of end users is done on client side. Acceptance to authorized user is done by TPA for transaction of data stored at CSP. Tools used for acceptance testing are P-unit,etc.

Output of Acceptance testing: The test is performed successfully by generating digital signature for the encrypted partitions of data.

Recognizing the need for a method of insuring the quality and correctness of AES implementation,SHA-3 implementation and MD5 implementation. NIST developed a set of tests and a testing protocol for the algorithms. It chartered independent testing to administer the tests and monitor the results. This set of tests is referred to as Validation. To achieve NIST certification a data security vendor must pass hundreds of different tests designed to validate that the vendor is properly encrypting and decrypting data for AES implementation. NIST testing is issued for AES Validation certificates.

X. CONCLUSION

XI. In this paper we have created a private cloud and secured the contents from the cloud server from making any changes in the contents of the client and this is done by using a third party auditor (TPA). The Third Party Auditor has performed storage auditing without the original copy of the data. The Cloud Service Provider stores all data of the client.The client creates its own Hash value. Then in the auditing section when the role of TPA comes, the TPA picks up a random file n generates a Hash value and then matches with the Hash value of the client, if the Hash value does not match then the data has been changed. In this we are using two algorithms SHA-3 and AES for encryption and decryption process.

After all this the system will achieve

1. Public auditability
2. Privacy Preserving
3. Correctness
4. Dynamics
5. Lightweight

For integrity of data

XII. REFERENCES

1. http://ieeexplore.ieee.org

Privacy preserving public auditing for secure cloud storage.

2. www.iosrjournals.org
Overview on security issue in cloud computing
3. www.iosrjournals.org/ccount
Requirements and challenges for securing cloud application and services.
4. www.ijarcsse.com
Using third party aditor for cloud data security:a review
5. ieeexplore.ieee.org
IT auditing to assure a secure cloud computing.
6. www.borjournals.com Efficient data sharing in cloud with third party auditor:a review study.
7. www.eprint.iacr.org
Public auditing for ensuring cloud data storage security with zero knowledge privacy.
8. www.cs.utu.fi/rlahdelma Introduction to cryptography.