Cloud Storage Security Using Encryption and Third-Party Storage Auditing Service

Rana M Pir
Lecturer
Leading university, Sylhet Bangladesh

***Abstract* - Cloud computing aims to enable end-users to easily create and use software without a need to worry about the technical implementations and the software's physical hosting location, hardware specifications, efficiency of data processing. Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage applications and services) .User can store that data remotely without maintaining local copy of data So the integrity of the data is major problem in cloud storage. Recently many works focuses on providing data dynamics and public verifiability for checking the remote integrity with the help of third party verifiers. Integrity in cloud computing is achieved by the Boneh-Lynn-Shacham (BLS) Signing algorithm that signing the data block before sending data to the cloud. To remove online burden of user we have public audit ability for cloud data storage. For that we use external audit party to check integrity of outsource data. We introduce third party auditor (TPA) for audit outsourced data without demanding local copy of user. No additional online burden for the cloud user and that can achieved by using Privacy-Preserving Public key based Homomorphic authenticator. Homomorphic authenticator efficiently supports public key based audit ability without having retrieval of the data blocks.**

***Index Terms* - Cloud Computing, Third Party auditor, Integrity verification, Signature verification, Metadata**

I. INTRODUCTION TO THIRD PARTY AUDITING IN CLOUD COMPUTING

Cloud computing is a general term for anything that involves delivering hosted services over the internet. These services are broadly divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS).

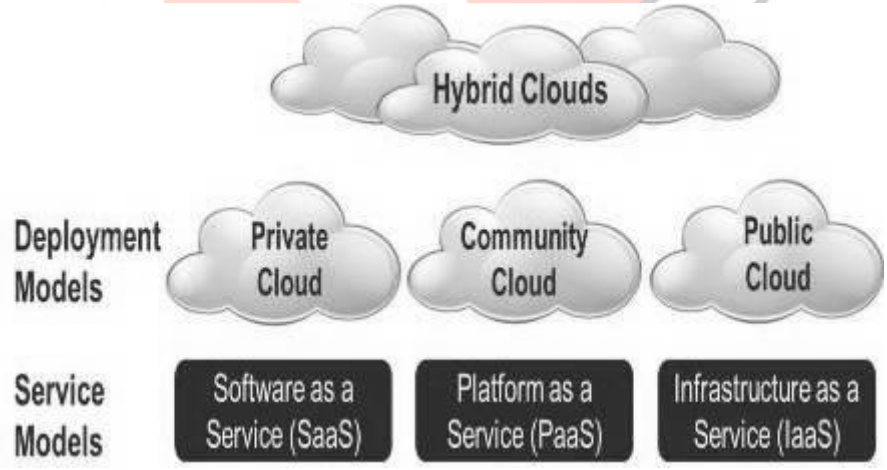

Fig 1

A cloud service has three distinct characteristics that differentiate it from traditional hosting. It is sold on demand, typically by the minute or the hour it is elastic a user can have as much or as little of a service as they want at any given time and the service is fully managed by the cloud service provider (the consumer needs nothing but a personal computer and Internet access).The advantage of cloud is cost saving. The prime disadvantage is security. Cloud computing is used by many software industries nova-days. Since the security is not provided in cloud, many companies adopt their unique security structure. Introducing a new and uniform security structure for all types of cloud since the data placed in the cloud is accessible to everyone, security is not guaranteed.

To ensure security, cryptographic techniques cannot be directly adopted. Sometimes the cloud service provider may hide the data corruptions to maintain the reputation. To avoid this problem, we introduce an effective Third party auditor (TPA) to audit the user's outsourced data when needed. The security is achieved by signing the data blocks. Signing is performed using BLS algorithm. We utilized public key based homomorphic authenticator with random masking to achieve privacy preserving auditing protocol. TPA performs the auditing task for each user i.e. single auditing. This increases the auditing time and computation overhead. The technique of Bilinear Aggregate Signature is used to achieve batch auditing and Data Dynamics (Insert, Delete, and Update).

Third Party Auditor

It supports an external auditor to audit the user's outsourced data without learning knowledge on the data content.

- Achieves batch auditing where multiple delegated auditing tasks from different users can be performed simultaneously by the TPA. Also supports dynamic operations on data blocks i.e. data update, append and delete.

Cloud computing components are classified as:

1. Cloud User (CU)
2. Cloud Service Provider (CSP) & Cloud Server (CS)
3. Third party Auditor (TPA)

Now let's get to know the component working for cloud computing in detail.

1. Cloud User (CU)
   Cloud user who has large amount of data files to be stored in the cloud. They may also dynamically interact with the CS (Cloud Server) to access and update their stored data for various application purposes.
2. Cloud Service Provider (CSP) or Cloud Server (CS)
   Services made available to users on demand via the Internet from a cloud computing provider's servers as opposed to being provided from a company's own on-premises servers. Cloud services are designed to provide easy, scalable access to applications, resources and services, and are fully managed by a cloud services provider. A cloud service can dynamically scale to meet the needs of its users, and because the service provider supplies the hardware and software necessary for the service, there's no need for a company to provision or deploy its own resources or allocate IT staff to manage the service. Examples of cloud services include online data storage and backup solutions, Web-based e-mail services, hosted office suites and document collaboration services, data- base processing, managed technical support services and more; and the person or authority who manages it is called as Cloud Service provider.
3. Third party Auditor (TPA)
   The third party auditor (TPA), who has expertise and capabilities that cloud users do not have and is trusted to assess the cloud storage service security on behalf of the user upon request. Users rely on the CS for cloud data storage and maintenance. They may also dynamically interact with the CS to access and update their stored data for various application purposes. The users may resort to TPA for ensuring the storage security of their outsourced data, while hoping to keep their data private from TPA. We consider the existence of a semi-trusted CS as does. Namely, in most of time it behaves properly and does not deviate from the prescribed protocol execution. However, during providing the cloud data storage based services, for their own benefits the CS might neglect to keep or deliberately delete rarely accessed data files which belong to ordinary cloud users. Moreover, the CS may decide o hide the data corruptions caused by server hacks or Byzantine failures to maintain reputation. We assume the TPA, who is in the business of auditing, is reliable and independent, and thus has no incentive to collude with either the CS or the users during the auditing process.

TPA Should be able to efficiently audit the cloud data storage without local copy of data and without bringing in additional on-line burden to Cloud Users.

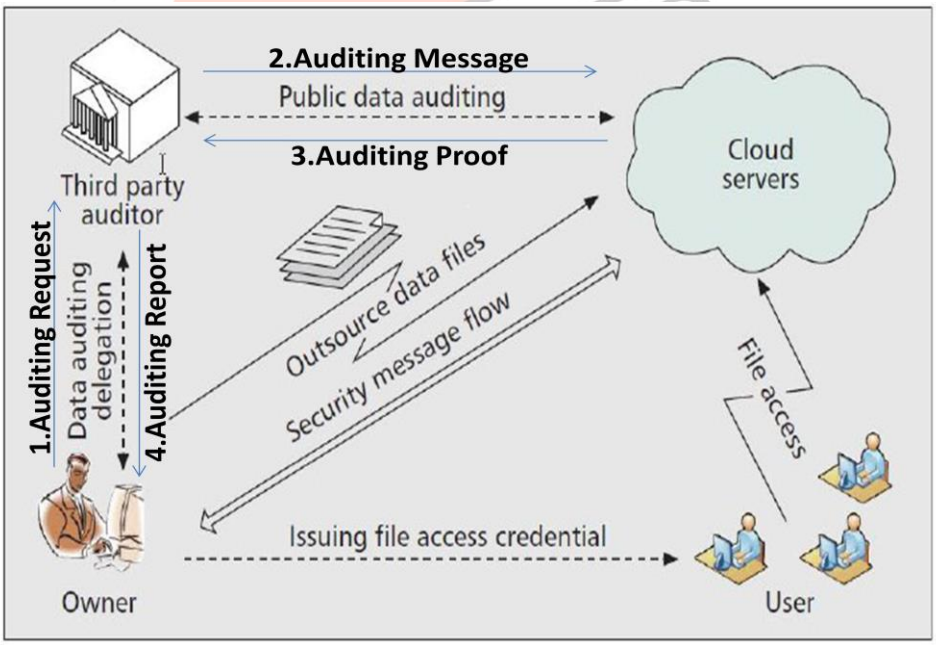

Fig 2 Architecture for Cloud Storage Devices [4]

II. METHODOLOGY AND EXPECTED OUTCOME

Before giving our main result, we first start with two warm up schemes. The first one does not ensure privacy-preserving guarantee and is not as lightweight as we would like. The second one overcomes the first one, but suffers from other undesirable systematic demerits for public auditing: bounded usage and auditor statefulness, which may pose additional on-line burden to users

as will be elaborated shortly. We believe the analysis of these basic schemes will lead us to our main result, which overcomes all these drawbacks.

Notation and Preliminaries

- F – the data file to be outsourced, denoted as a sequence of n blocks m1,...,mn ∈ Zp for some large prime p.
- MAC (·) (·) – message authentication code (MAC) function, defined as: K × {0,1}∗ → {0,1}l where K denotes the key space.
- H (·), h (·) – cryptographic hash functions. We now introduce some necessary cryptographic background for our proposed scheme.

Bilinear Map. Let G1, G2 and GT be multiplicative cyclic groups of prime order p. Let g1 and g2 be generators of G1 and G2, respectively. A bilinear map is a map e: G1 × G2 → GT such that for all u ∈ G1, v ∈ G2 and a, b ∈ Zp, e (u^a,v^b) = e(u,v)$^{ab.}$ This bilinearity implies that for any u1,u2 ∈ G1, v ∈ G2, e(u1 · u2,v) = e(u1,v) · e(u2,v). Of course, there exists an efficiently computable algorithm for computing e and the map should be non-trivial, i.e., e is non- degenerate: e (g1, g2) ≠ 1.

Basic Scheme I: The cloud user pre computes MACs σi = MACsk(i||mi) of each block mi (i ∈ {1, . . . , n}), sends both the data file F and the MACs {σi}1≤i≤n onto the cloud server, and releases the secret key sk to TPA. During the Audit phase, the TPA requests from the cloud server a number of randomly selected blocks and their corresponding MACs to verify the correctness of the data file. The insight behind this approach is that auditing most of the file is much easier than the whole of it. However, this simple solution suffers from the following severe drawbacks:

1) The audit from TPA demands retrieval of users' data, which should be prohibitive because it violates the privacy-preserving guarantee

2) Its communication and computation complexity are both linear with respect to the sampled data size, which may result in large communication overhead and time delay, especially when the bandwidth available between the TPA and the cloud server is limited.

Basic Scheme II: To avoid retrieving data from the cloud server, one may improve the above solution as follows: Before data outsourcing, the cloud user chooses s random message authentication code keys {skt} 1≤ t≤s, pre-computes s MACs, {MACskt(F)}1≤t≤s for the whole data file F, and publishes these verification metadata to TPA. The TPA can each time reveal a secret key skt to the cloud server and ask for a fresh keyed MAC for comparison, thus achieving privacy-preserving auditing. However, in this method:

1) The number of times a particular data file can be audited is limited by the number of secret keys that must be a fixed priori. Once all possible secret keys are exhausted, cloud user then has to retrieve data from the server in order to re-compute and re-publish new MACs to TPA.

2) The TPA has to maintain and update state between audits, i.e., keep a track on the possessed MAC keys. Considering the potentially large number of audit delegations from multiple users, maintaining such states for TPA can be difficult and error prone.

Note that another common drawback of the above basic schemes is that they can only support the case of static data, and none of them can deal with data dynamics. For the reason of brevity and clarity, our main result will focus on the static data, too.

Implementation Schemes:

In my Implementation Scheme having three components:

i) Cloud User (CU)

ii) Cloud Service Provider (CSP) & Cloud Server (CS)

iii) Third party Auditor (TPA)

Privacy-Preserving Public Auditing Scheme:

I use the technique to uniquely integrate the homomorphic authenticator with random masking technique. In my system, the linear combination of sampled blocks in the server's response is masked with randomness generated by a pseudo random function (PRF).With random masking, the TPA no longer has all the necessary information to build up a correct group of linear equations and therefore cannot derive the user's data content, no matter how many linear combinations of the same set of file blocks can be collected. Meanwhile, due to the algebraic property of the homomorphic authenticator, the correctness validation of the block-authenticator pairs will not be affected by the randomness generated from a PRF. A public auditing scheme consists of four algorithms (KeyGen, SigGen, GenProof, and VerifyProof). KeyGen is a key generation algorithm that is run by the user to setup the scheme.

SigGen is used by the user to generate verification metadata, which may consist of MAC, signatures, or other related information that will be used for auditing. GenProof is run by the loud server to generate a proof of data storage correctness, while, VerifyProof is run by the TPA to audit the proof from the cloud server. I propose a batch Auditing scheme based on the BLS signature algorithm.

Implementation BLS signature scheme:

The signature scheme is then given as the set of algorithms

{KeyGen; SigGen; Verify}

Algorithm 1: generates an asymmetric key pair (x, v) ∈ Zn * Gn with private key x and public key v.

KeyGen

Data: generator g2 for G2, prime number p
Result: private key x ∈ Zn, public key v ∈ G2
Choose random x ∈ Zn
V ⟵ $g2^x$
Return (x, v)

Algorithm 2: This used when signing a message M with the private key x.This algorithm requires a hash function H that can hash the message to an element h ∈ G1. We will assume that H is a random hash function.

SigGen

Data: private key x ∈ Zn, message M ∈ {0,1}*
Result: signature σ ∈ G1
h ⟵ H(M) ∈ G1
σ ⟵ h^x
Return σ

Algorithm 3: Verify the signature with public key.

Verify

Data: public key v ∈ G2, message M ∈ {0,1}*, signature σ ∈ G1
Result: boolean value
h ⟵ H(M) ∈ G1
Return Test ((g2, v, h, σ)

Proposed Scheme Implementation scenario:

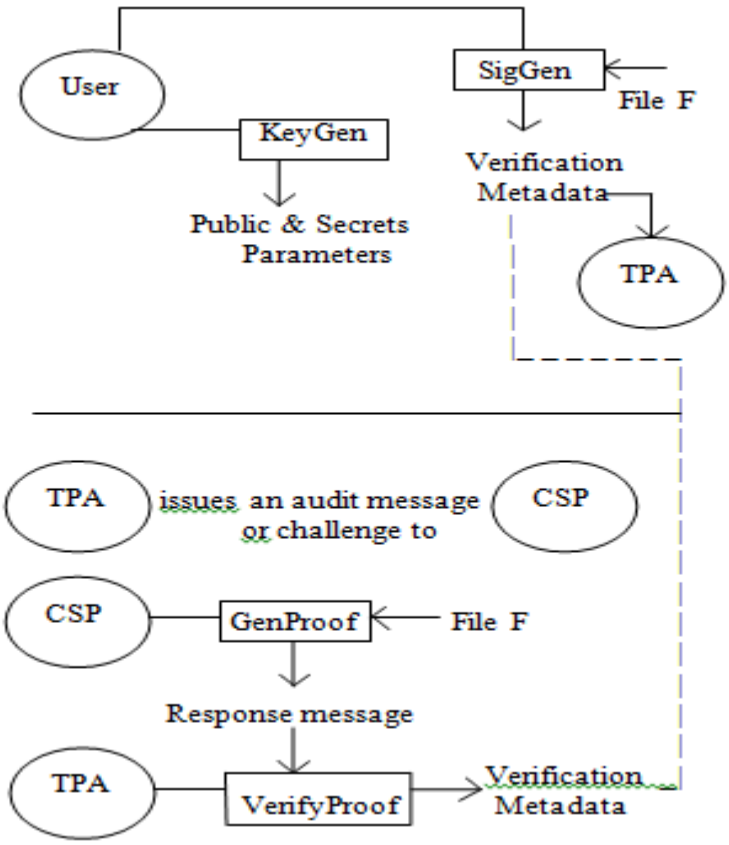

Fig 3 Proposed Scheme Implementation scenario

III. IMPLEMENTATION FLOW

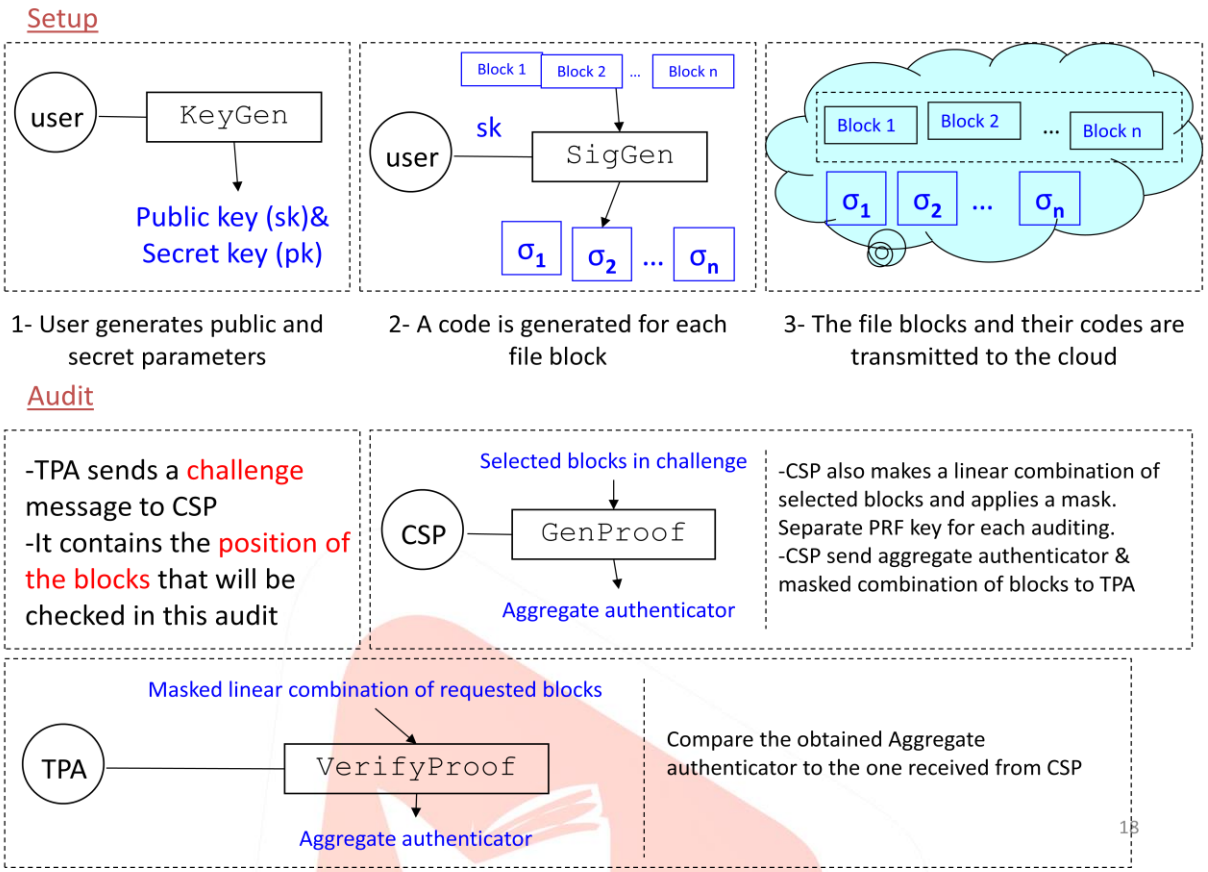

Fig 4 Implementation Flow (a)

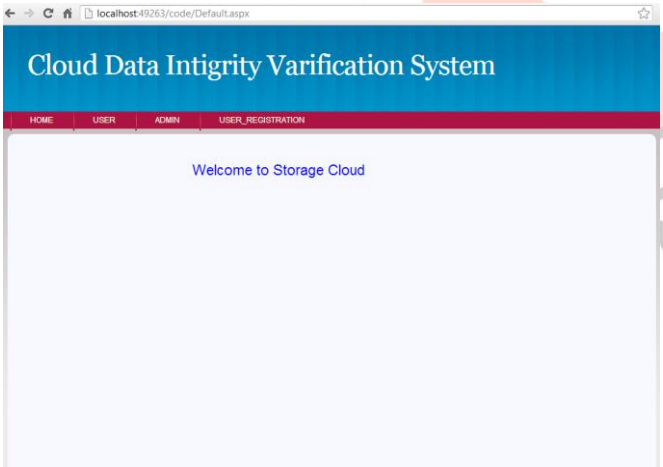

Fig 5 Implementation Flow (b)

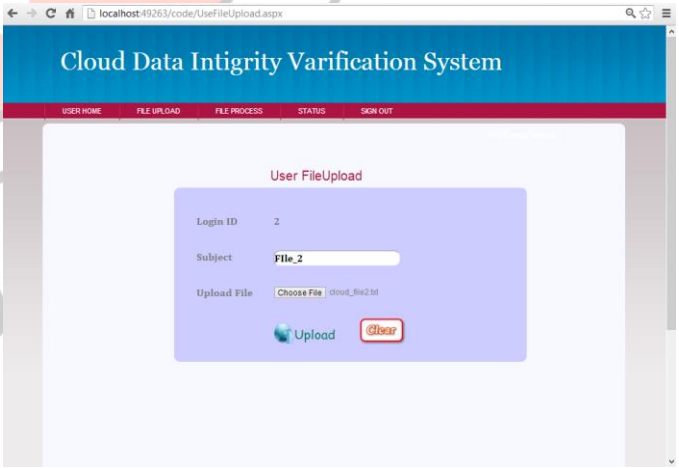

Fig 5 Implementation Flow (b)

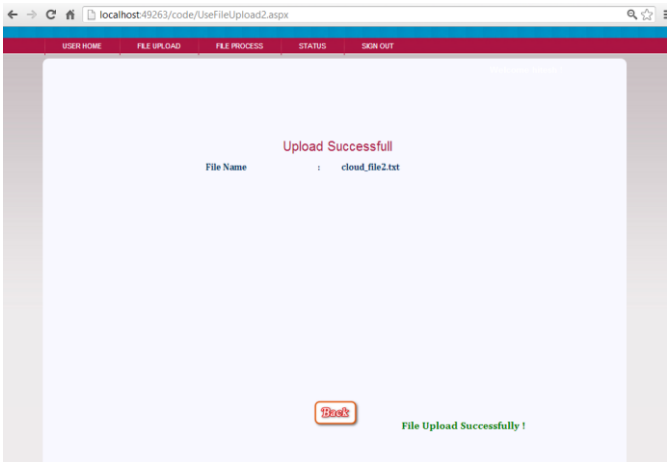

Fig 6 Implementation Flow (c)

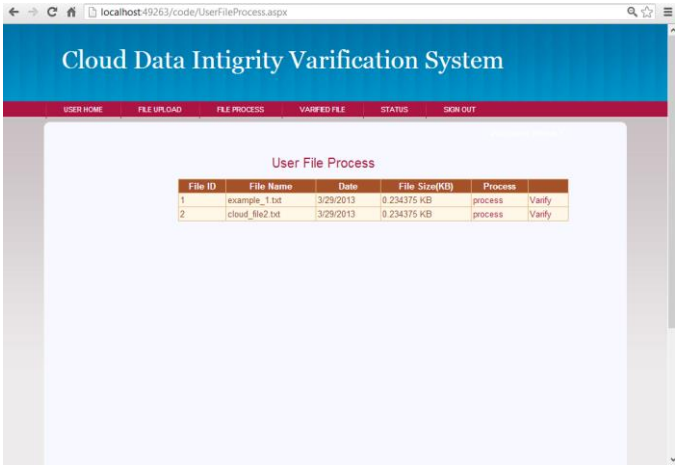

Fig 7 Implementation Flow (d)

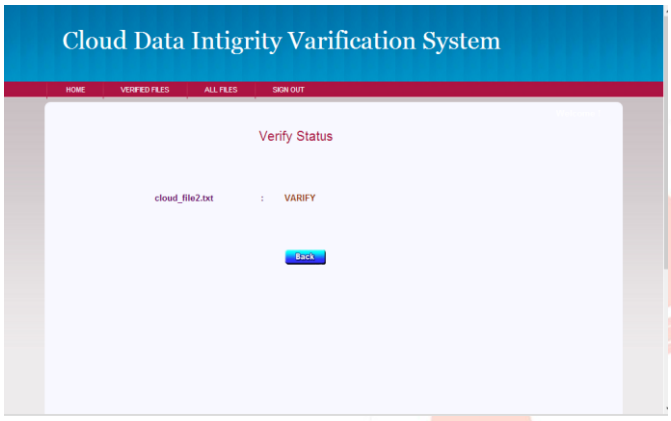

Fig 8 Implementation Flow (e)

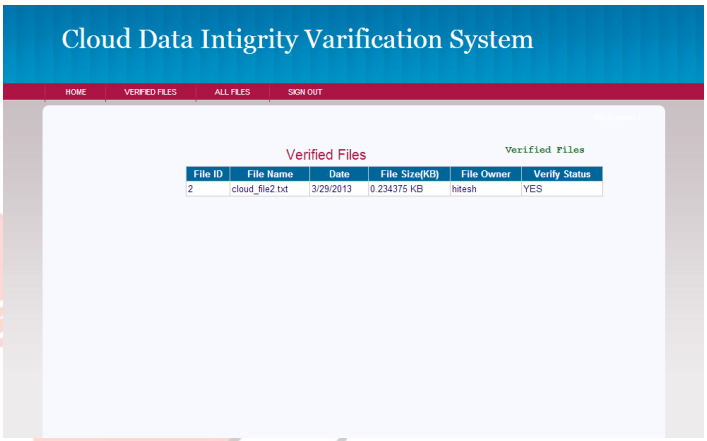

Fig 9 Implementation Flow (f)

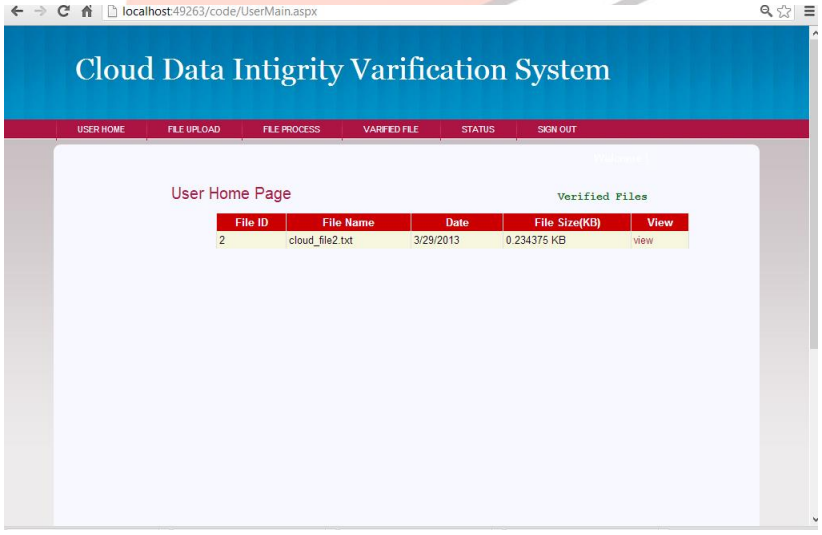

Fig 10 Implementation Flow (g)

IV. CONCLUSION

In privacy-preserving public auditing system for data storage security in cloud computing utilize the homomorphic linear authenticator and random masking to guarantee that the TPA never know the data content which is stored on cloud server during the efficient auditing phase .it's also support batch auditing and data dynamics (insert, update, delete).

REFERENCES

[1] Cloud Security Alliance, "Security Guidance for critical areas of focus in Cloud Computing V3.0" https://cloudsecurityalliance.org/guidance/csaguide.v3.0.pdf

[2] National Institute of Standards and Technology- Computer Security Resource Center www.csrc.nist.gov
[3] http://en.wikipedia.org/wiki/Cloud_computing
[4] Jachak K.B.*,Korde S.K.,Ghorpade P.P. and Gagare G.J. "Homomorphic Authentication with Random Masking Technique Ensuring Privacy & Security in Cloud Computing, BIOINFO Security Informatics -2012
[5] Cong Wang, Student Member, IEEE, Sherman S.-M. Chow, Qian Wang, Student Member, IEEE,Kui Ren, Member, IEEE, and Wenjing Lou, Member, IEEE "Privacy-Preserving Public Auditing for Secure Cloud Storage, IEEE-2012
[6] Cong Wang1, Qian Wang1, Kui Ren1, and Wenjing Lou2," Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing", IEEE INFOCOM 2010, San Diego, CA, March 2010
[7] Ms. Vaishnavi Moorthy1, Dr. S. Sivasubramaniam2," Implementing Remote Data Integrity Checking Protocol for Secured Storage Services with Data Dynamics and Public Verifiability In Cloud Computing, IOSR Journal of EngineeringMar. 2012, Vol. 2(3) pp: 496-500
[8] Caroline Fontaine and Fabien Galand,"A Survey of Homomorphic Encryption for Nonspecialists",EURASIP Journal on Information Security, pages 1 to15, January 2007.
[9] Rosario Gennaro and Daniel Wichs, Fully Homomorphic Message Authenticators , IBM Research, T.J. Watson, May 23, 2012
[10] Nitin Jain, Saibal K. Pal & Dhananjay K. Upadhyay." Implementation And Analysis Of Homomorphic Encryption Schemes" International Journal on Cryptography and Information Security (IJCIS), Vol.2, No.2, June 2012
[11] Kerckhoffs, A., (1883). "La cryptographie militaire (part i)", Journal des Sciences Militaires, Vol. 9, no. 1, pp. 5–38.
[12] Kerckhoffs, A., (1883). "La cryptographie militaire (part ii)", Journal des Sciences Militaires, Vol. 9, no. 2, pp. 161–191.
[13] http://en.wikipedia.org/wiki
[14] H. Shacham and B. Waters, "Compact proofs of retrievability," in Proc. of Asiacrypt 2008, vol. 5350, Dec 2008, pp. 90–107.
[15] 104th United States Congress, "Health Insurance Portability and Accountability Act of 1996 (HIPPA)," Online at http:// asp.ehhs.gov/admnsimp/pl104191.htm, 1996.
[16] Amazon.com, "Amazon s3 availability event: July 20, 2008,"Online at http://status.aws.amazon.com/s3-20080720.html,2008.
[17] D. Boneh, B. Lynn, and H. Shacham, "Short signatures from the Weil pairing," J. Cryptology, vol. 17, no. 4, pp. 297–319,2004.