Public Auditability and Privacy preserving in Cloud Storage

Kishan Lathkar, Ambulgekar H. P
SGGS IE&T
Nanded, Maharashtra
India
lathkarkishan@sggs.ac.in, ambulgekar@sggs.ac.in

Abstract: *Cloud computing is a revolutionary new approach to how computing services are produced and consumed. It is a concept of sharing resources and giving them as essential resources. Using cloud computing resources, data, computations, and services can be shared over scalable network of nodes; these nodes may represent the datacenters, cloud users and web services. On the same note cloud storage talk about storing the data on a remote storage located at other organization's infrastructure. The data storage is maintained and managed by the organization; the user will pay for the storage space which is used. Outsourcing data ultimately relinquishes the control of data from user and the lot of data is in control of the cloud server. The cloud server is succeeded by cloud service provider which is a different administrative entity, so ensuring the data integrity is of prime importance. This paper studies the problems of ensuring data storage correctness and proposes an effective and secure scheme to address these issues. A third party auditor (TPA) is introduced securely. Who will on behalf of users request will periodically verify the data integrity of the data stored on cloud server. There will not be any online burden on user and security of data will be maintained as the data will not be shared directly with the third party auditor. A homomorphic encryption scheme is used to encrypt the data by using Elliptic curve Digital Signature Algorithm (ECDSA) which will be shared with the TPA. ECDSA provides efficient and secure solutions for the cloud storage servers. It leads to fast computation time, reducing in processing power, save the storage and bandwidth. The results can be further extended to enable the third party auditor to do multiple auditing simultaneously*

Keywords: Cloud Storage, Data Integrity, Privacy Preserving, Public Auditability, Cloud Service Provider (CSP), Third Party Auditor (TPA)

Recieved 30 October 2014, Revised 2 December 2014, Accepted 9 December 2014

I. Introduction

Cloud computing has grew a lot of attention in recent years as Cloud computing is characterized as a style of computing empowered the abilities delivered are "as a service" to client that utilizing the internet technology. The cloud computing deals with few benefits like pay-for- use, lower costs, fast deployment, versatility, quick provisioning, rapid flexibility, universal network access and data storage solution, on-demand security controls, real time detection of system tampering and fast re-constitution of services. Cloud computing basically stores all of cloud computing applications and databases in the data center which are put at distinctive areas. Because of this development of software programming, data and service are

are not trustworthy. So this provide upsurge to several security challenges like: virtualization vulnerabilities, availability vulnerabilities, web application vulnerabilities, privacy and control issues emerging from outsiders having physical control of data, issues identified with identity and credential management, data verification, data modification, integrity, privacy, data loss are issues recognized with authentication of the respondent device to devices and IP spoofing.

Cloud computing is considered as the revolutionary technology due to the advantages it provides to its users like usage based assessing, location independent pooling of resources, scalable resources, maintenance and upgrade is simplified, on demand service, broad network access. These advantages make cloud computing very appealing but it brings in new challenges and threats towards the security of data stored on cloud server. The data is stored on cloud server which is hosted in other organization's infrastructure and is managed by the cloud service provider. Amazon's Elastic Compute Cloud (EC2) and Amazon Simple Storage Service (S3) and apple icloud are examples of cloud storage. As the data is stored on cloud server, the risk of data integrity or the storage correctness of the data is increased.

As the data is stored on cloud storage, the risk of data integrity or the storage correctness of the data is increased. The infrastructures defined for cloud are very powerful and reliable than other personalized computing devices still they face wide range of internal and external security threats. There are several instances of security breaches of cloud services which are being reported. The cloud service provider (CSP) for regulatory reasons may delete the data which is not accessed or is infrequently accessed or the CSP may hide the data loss incidents on the way to maintain status. Though cloud storage gives many advantages to users to store their data on cloud server but CSP does not give guarantee the data integrity of the outsourced data. There are two very simple ways in which the user can verify the data integrity, first is the user can download all the data and verify the integrity of the outsourced data. The solution is impractical as the I/O operations over network are expensive. Second solution is user can check the storage correctness of the data whenever that data is accesses but this solution does not give the assurance of data integrity of the data which is not accessed or is rarely accessed. The confidentiality issue of outsourced data can be handled by using encryption but it is very difficult to verify the data integrity of the data without having the local copy of the data. As discussed above downloading the data for verifying data integrity is not an efficient way.

To gain trust of users on cloud storage, the data need to be securely stored on cloud and strong assurance should be given to users that their data is not tampered or deleted by cloud service provider (CSP). More over users should be able to use the data without doing too many operations in addition to downloading the data. The task of auditing the data to ensure data integrity in cloud is expensive and hectic for users, also the user wants to avoid the complications of computation and communication of verifying of data integrity. Apart from single user there will be many users using cloud storage so it will be easier to manage integrity verification request from single designated entity instead of getting request from multiple users over the network.

We are proposing a system which securely introduce privacy-preserving third party auditing (TPA) in cloud storage system. To guarantee the data integrity of outsourced user's data, save user's computation time and online burden of additional processing. TPA will verify the storage correctness of the outsourced data periodically when user will initiate request for verification. The TPA will have the ability and resources to do the auditing process for verifying data integrity and will help to reduce burden of users and save their computation resources. Enabling public auditing will not raise any security concerns for user's outsourced data. More over the evaluation done by TPA will also support to the cloud service provider (CSP) to improve their cloud services and to gain belief in cloud storage service.

1.1 Overview Of Paper

In the next section we examine the privacy in cloud storage. From this we define the important privacy issue related to cloud storage. In a problem statement we introduce the system model and general architecture of proposed system. Also define the design goals of propose work. Then in the next section we give an overview of the privacy preserving public auditing scheme and overview of proposed algorithm. We look at key features of mathematical model of Elliptic curve digital signature algorithm and Markel hash tree. We follow by reviewing related work on privacy preserving and third party auditing. Finally we finish with our conclusions and list further work that remains to be completed.

2. Privacy in Cloud Storage

A typical way to supply data subject for privacy of data storage on cloud is data and control approach [8]. Legally we store the data on cloud and other computing applications have level of privacy security which relies on type of data and cloud cloud and other computing applications have level of privacy security which relies on type of data and cloud server. For privacy preserving

in distributed computing applications which mainly focus on authentication and authorization technology connected to policy. These procedures based on when implementation is done at the point where data is accessed. The cloud makes data implementation is difficult at this level. In the cloud the data can be split and stored at different cloud server. It is difficult to recreated the services for provider and for identify the trust across the field. In addition to this, the data is moved quickly and take policy decision at the specific points of the data in such case of data is unrealistic.

The stored data can be accessed by cloud infrastructure provider, which also accessed this data at any non-application specific level. The cloud is utilized for vanilla operation, for example, storage or transforming of data. Data can be encrypted before send to the cloud server and next to no logical data can be connected with the data when compared with data stored at the service provider level near to the application logic. Subsequently the cloud can be seen as the layer down from the service provider, and this permits privacy to be taken care of at a more unique level from direct policy.

2.1 Definition of Privacy

In privacy human can be free from all privacy control and interferences, so it can maintain degree of intimacy. If privacy can break it may create problem to the cloud user. Privacy provides the security for the truthful utilization of data for cloud user. The American Institute of Certified Public Accountants (AICPA) and Canadian Institute of Charted Accountants (CICA) stated that, "*Privacy is the right and obligation of individuals and organizations with respect to the collection, use, retention, and disclosure of personal information*".

2.2 Privacy Issues in Cloud Storage

The most important objective in cloud storage is security which securing the data privacy. But as we discuss about earlier issues, it is hard to prevent threats in cloud storage because it is relay on storage system. The data will be presented but there is risk of unauthorized access. There is big challenge to securing user privacy in cloud storage. The essential part of this challenge is data segregation where users data is schedule from others user's data.

Another difficult problem of cloud storage is development of data. Data secrecy is solution for data privacy and security. The data privacy of user is difficult when it contain user's personal information and data in the cloud and also with development of cloud storage. The privacy of users data access in cloud computing is still in progress and still more user's objectives.

3. Problem Statement

3.1 System Model

"*To securely introduce a Third Party Auditor for verifying data integrity of outsourced data in a cloud storage system with the privacy preserving property* ", General architecture of the cloud storage system is shown in Figure 1.The proposed system has three entities as explained as follows:

1. Cloud User: upload the data files on cloud server along with the verification metadata. The files can be downloaded whenever user wants to modify the file. The metadata will be modified for the modified data file. User will initiate the key generation.

2.Cloud Server: It has the storage space required for storing large number of data files. This one is achieved by the cloud service provider (CSP). Cloud server will be challenged by the TPA to which it respond by generating proof of integrity.

3.Third Party Auditor: TPA will perform the auditing process on user's request. For auditing it will get the verification metadata from the user and based on that it will send challenge to the cloud service provider. The reply sent by the cloud server will then be verified by the TPA.

3.2 Design Goals

Implementation of the system gives the following performance guarantee.

1. Storage Correctness- Without retrieving the whole data from cloud server TPA will verify the storage correctness.

2. Support Auditing-Cloud server will not be able to pass the auditing process without storing the data intact.

3. Privacy-Preserving-Auditing process will be privacy preserving, That is TPA may not be able to draw the original data from the information available during auditing.

4. **Efficient**-TPA will perform the auditing process with minimum computation and communication.

5. **Multiple Auditing**-TPA will deal with multiple auditing requests for multiple users.

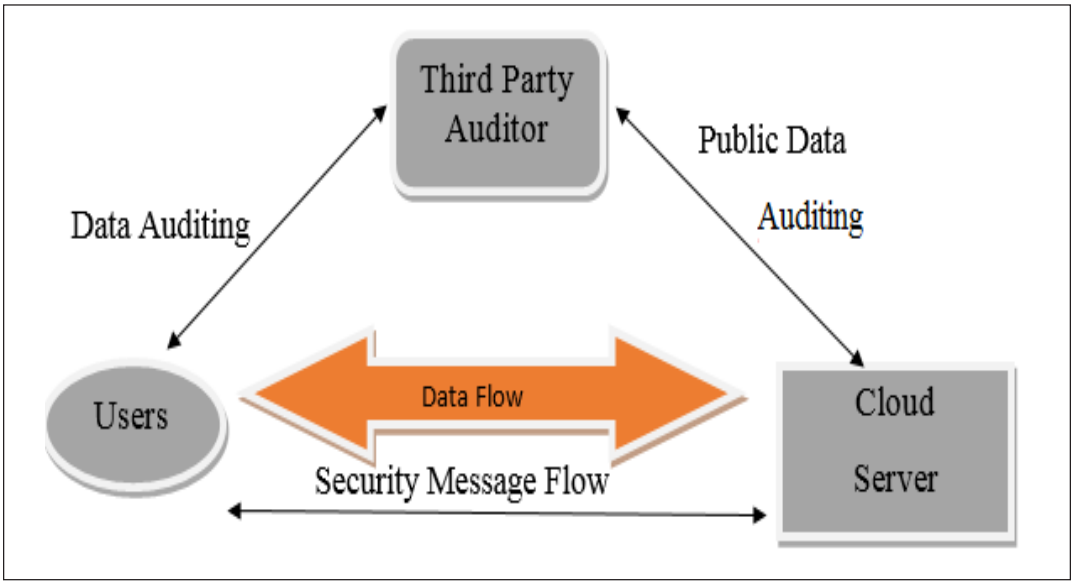

Figure 1. General architecture of cloud storage

4. Proposed Work

A privacy preserving public auditing scheme will be made up of four algorithms,

1. Key-Generate: The user will initiate this algorithm for generating keys.

2. Sig-Generate: By using this algorithm the user can generate verification metadata.

3. Generate-Proof: The proof of data integrity by the cloud server is generated by using this algorithm.

4. Verify-Proof: TPA uses this algorithm to audit the proof generated by the cloud server.

The Public auditing scheme will be made up of two phases, Setup phase and Audit phase

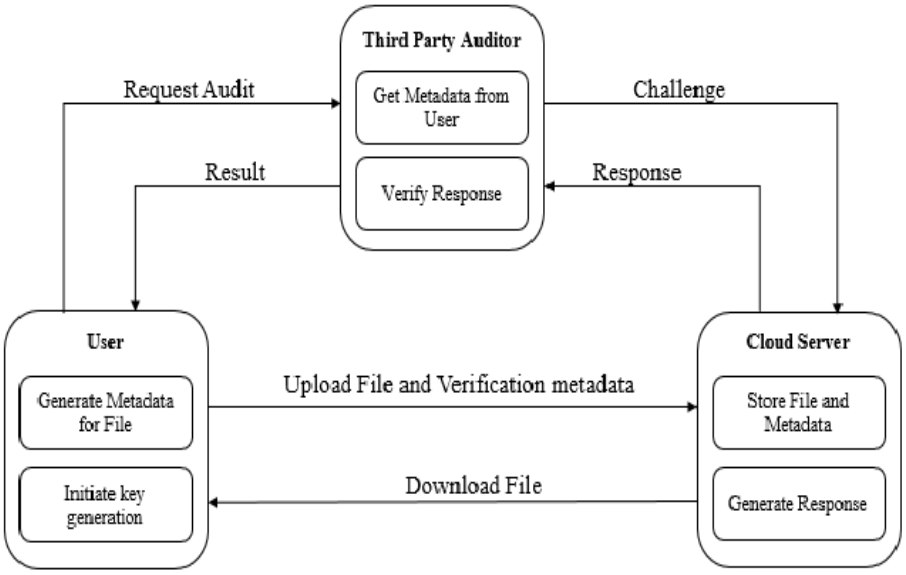

Figure 2. System Architecture.

4.1 Setup Phase

1. The cloud user will initialize the generation of the keys (public and private) by executing KeyGen ( ).

2. Verification metadata will be generated by processing the data file using SigGen ( ).

3. After generating metadata cloud user will store the encrypted data file and metadata on cloud server and remove the data file from local storage.

4. Verification metadata also send to Third Party Auditor (TPA) to verify the integrity of data file stored on a cloud server.

4.2 Audit Phase

1. The TPA will send a challenge message to cloud server for checking the data integrity.

2. Cloud server will execute GenProof ( ) by using the metadata to generate the proof of data integrity. Then Send generated proof to the TPA.

3. TPA will then run VerifyProof ( ) to verify the response sent by the cloud server.

4.3 Mathematical Model

4.3.1 Elliptic Curve Digital Signature Algorithm

Elliptic Curve Cryptography [12] is a public key cryptography algorithm. Elliptic curve cryptography provides efficient and secure solutions for the cloud storage servers. It requires fewer bits than the conventional encryption technologies for provides the similar amount of security. It provides data integrity, confidentiality and data origin authentication. Compare with existing cryptosystem it must be a smaller key size, fast computation time, saves the storage and bandwidth, reduce the processing power. In public auditability scheme ECC signature algorithm mainly consists of four phases. These are key generation, signature generation, Generation proof and signature verification. The signature generation algorithm use the user's private key to generate the signature. The Generation proof algorithm use the user's public key to generate the signature at server.

The ECDSA algorithm need the following domain parameters for perform Key generation, signature generation and Generation Proof, The following are notations and preliminaries, Domain parameters $D = (p, q, Zq , b, G, n, h)$

1. Let q be a large prime of finite Field Zq called prime field.

2. If $q = p \wedge m$, where p is a prime and it is called characteristic of Zq and m is the extension degree of Zq, odd prime ($q = p$) or a power of 2 ($q = 2 \wedge m$)

3. Elliptic curve parameters a and b in Zq, which defines the equation of the elliptic curve E *over* Zq.

4. A bit string Seed s of length $\geq$ 160 bits (Optional).

5. The field point $G = (Ag, Bg)$ on curve E (Zq) i.e. $Ag,Bg \in Zq$.

6. The Order n of the point G: n is prime, with $n > 2\wedge160$ and $n > 4\sqrt{q}$

7. The Cofactor h: The order of the E (Zq) divided by the Order n of h.

4.3.1.1 KeyGen( ) :

To generate the key pair is related with a particular set of elliptic curve domain parameters $D = (p, q, Zq, a, b, G, n, h)$, E is an elliptic curve defined over finite field Zq, and G is a point of prime order n in $E(Zq)$, A field size q, where q is odd prime. The user should perform the following,

1. Choose any random integer Xa in the interval [1, n–1].

2. Compute $Ya = Xa * G$.

3. A key pair is (Xa,Ya) where Ya is public key and Xa is Private key.

4.3.1.2 SigGen( ) :

1. Compute $\gamma = H(m)$, where *H* is a cryptographic hash function, such as SHA-1.

2. Select a random integer *k* form, $1 \leq k \leq n-1$.

3. Compute $\alpha = Ag \bmod n$, where $(Ag, Bg) = k * G$, If reminder $\alpha = 0$; then go to step 2.

4. Compute $k \wedge -1 \bmod n$.

5. Compute $\beta = k\wedge -1 (\gamma + Xa * \alpha) \bmod n$, If $\beta = 0$ then go to step 2.

6. The signature Pair is (α, β).

4.3.1.3 GenProof( ):

GenProof algorithm is run by the cloud server to generate a proof of data storage correctness. The message *m* with respect to (α,β). Cloud service provider gets an authenticated copy of cloud user domain parameters *D* =(*p, q, Zq, a, b, G, n, h*) and the public key *Ya* then do the following.

1. $\sigma = H(m)$, where H is the same hash function used in the signature generation at user side.

2. Compute $w = \beta \wedge -1 \bmod n$.

3. Compute $u1 = \sigma w \bmod n$ and $u2 = áw \bmod n$.

4. Compute $(Ag,Bg) = u1G + u2Ya$.

5. Compute $v = Ag \bmod n$.

6. Cloud server send *v* as the response proof of storage correctness to the TPA.

4.3.1.4 VerifyProof ( ) :

With the response, The TPA then verifies the response via VerifyProof. Accept the signature if andonly if $v = á$.

4.4 Merkle Hash Tree

Figure 3 shows the Merkle hash tree (MHT). [5] It is constructed as a binary tree. The MHT divides the parent node up to sub-blocks. The hash value is associated with every non leaf node. It is an authenticated structure it is proved that the set of elements are unaltered and not damaged. The MHT is used in the proposed Scheme to divides the user's file in to blocks.

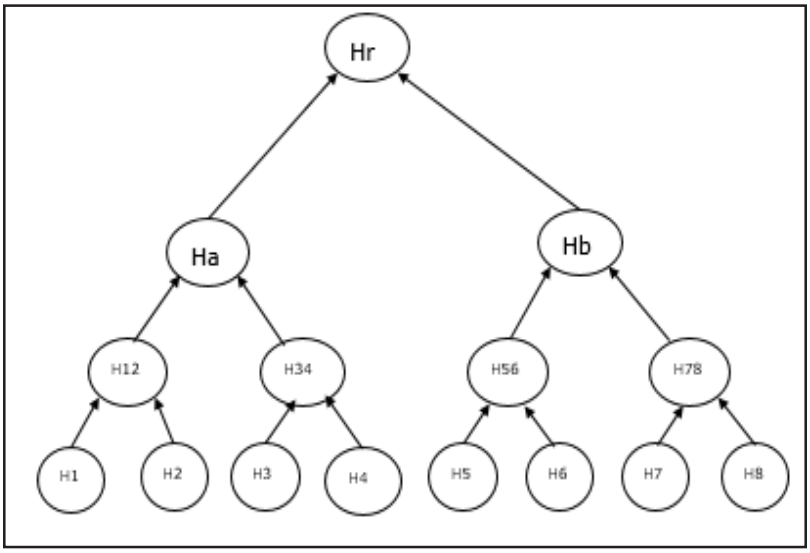

Figure 3. Merkle tree

age server, it is divided in to blocks by using MHT and the hash value will be allocated to each block. The blocks are read from the left to right sequence. The storage server stores the block number, the content associated with that block, and its calculated hash value.

4.5 Batch Auditing

The TPA is able to do batch auditing which will decrease the time of response for audit request from multiple users. The comparison of single and batch auditing for number of files that time required for batch auditing is less than the time required for auditing number of files one by one. By using batch auditing reduce communication and computation overhead.

5. Related Work

There are three ways for security services that connected to extensive range of data locally stored or stored on cloud server. Those services are used for data integrity, privacy and accessibility. The security services used for data stored remotely same before some complication arises. Cloud service provider provides secure data access to the user. There is creating system such that it is measurable to check the reliability of data which stored on cloud server. To solve the problem of data integrity in the cloud storage there exist many methods.

5.1 Encryption Method

There are some methods are available that makes utilization of encryption method that achieve the privacy in the cloud. Ru Wei et.al [6] proposed that privacy preserving in cloud storage used to solve the issues like data formation, data organization structure, keys management, user's communication and also support dynamic operation on data. It uses interactive protocol and key based induction algorithm. It gives assurance of data privacy, solve the problem of ineffectiveness key production, minimize the load of encryption and decryption, manage the number of keys, save the user's cloud storage space, minimize the execution time of system and perfectly provides privacy and security to different user's, data owner and cloud service provider. There is need of new technique to minimize the owner encryption burden and work on cipher text. There is problem related to the encryption based method that limits the data usage and puts additional burden. For reducing burden overhead there is access control mechanism available.

5.2 Scalable and Efficient Provable Data Possession[6]

There is another technique are available for integrity verification which is Provable Data Possession (PDP). The PDP introduced by Ateniese et.al. [6], [9] for ensuring keeping the data files on untrusted storage. For provable security it gets RSA-based homomorphic encryption techniques which get block less verification. This scheme depends upon symmetric key encryption and cryptographic hash function. The linear combination of file block is used for verification with no data privacy. The drawback of PDP is it checks Static data possession but not checks the dynamic data possession. Erway et.al. Was proposed the more advance strategy called Dynamic Provable Data Possession (DPDP). This was first method to support dynamic data possession. It was an extension of PDP model. However, this current method efficiency remain in questionable situation i.e. it doesn't completely public verifiability.

5.3 PORs: Proofs of Retrievability for Large Files[4]

Proof of Retrievability (POR) is technique implemented as protocol by Jules et.al. [4]. the users data must be stored at any remote storage server, this data not be unmodified and unaltered. The disadvantage of POR is doesn't shield the data from modification and that action performed by the cloud service provider. An improved POR technique was proposed by Shacham and Waters [8] consist of the use of BLS signatures. Homomorphic linear authenticators are used to build from secure BLS signatures which are publicly verifiable. Their approach is not suitable for privacy preserving.

5.4 Auditing to Keep Online Storage Services Honest[11]

M.A. Shah et al. [11] introduced a TPA check the honesty of online storage by first encrypt the data file to and stored it, then number of precomputed symmetric key used over the encrypted file and this file send to the auditor. The auditor role is to make sure of data file integrity and previously committed key held at server. The TPA scheme needs to maintain state and suffer from bounded usage which results in burden to the consumer when all symmetric key are hashed. This scheme works only for data files which are encrypted.

5.5 Towards Secure and Dependable Storage Services in Cloud Computing[7]

Wang et al. [2], [5] propose an effective and flexible scheme used for dynamic data support for data integrity of user data in cloud server. They use correcting coding to avoid the redundancies in the cloud and increase data dependencies, overcome communication and storage overhead. The code data ensures the storage correctness and data error localization which use token scheme for precomputation. With no explicit knowledge of data verification is done. The further work explores to maintain dynamic operation on data blocks like insert, delete and update on cloud storage data. The TPA is mainly used to support to check the integrity of data on cloud server and freely. The security and performance of this scheme is efficient and robust as compared to Byzantine failure, malicious data modification attack, and even server colluding attacks.-

5.6 Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing[5]

One more public auditing scheme propose by Wang and Lou et al.[5] to ensure cloud data privacy and security. This scheme is mainly motivated to the public auditing system for data storage privacy in cloud. It fully support to the fully dynamic data operation for block insertion. The task of TPA is trustworthiness to the cloud user for dynamic data operation. The TPA contribution to the cloud user is his/her data stored in cloud is surely in place, it can be vital accomplish to economy of scale in cloud computing. The most general data type of operation done by dynamic data operation for example, data modification, insertion and deletion. Further extension to this scheme is multiple batch auditing and task can be done by the TPA to the users. To achieve efficient data dynamics, they can expand the current proof of storage models by using Merkle Hash Tree (MHT) construction for block tag authentication.

5.7 Privacy-Preserving Public Auditing for Secure Cloud Storage[2]

In this paper C. Wang and Q. Wang et al. [1][2]proposed a third party auditing system for cloud based data storages whilst preserving privacy of the user's data. Homomorphic linear authenticator (HLA) and random masking has been utilized to guarantee that the TPA wouldn't be able to gain access to the data content stored on the cloud server not compromising the efficiency of the auditing process. This removes the burden of auditing task from the cloud user and also reduces the users' fear that their outsourced data may be accessed by a mistrustful party. The data file which is having linear combination of sampled blocks is masked with randomness generated by the server. This linear combination of blocks is given by the server's response. In random masking technique the block has chosen randomly. Further Extension of this scheme batch auditing TPA handles multiple users during auditing process. TPA performs auditing process with minimum communication.

6. Conclusion

In this paper highlights some privacy methods for overcoming the issues in privacy on untrusted data storage in cloud server. Also provide some solutions for the preserve privacy in secure cloud storage. Privacy security is a key issue for cloud storage and is to be considered critical. To guarantee that the risk of privacy has been alleviated a number of systems that may be utilized in order to achieve privacy. We propose an efficient and inherently secure homomorphic linear authenticator scheme which is based on Elliptic Curve Digital Signature Algorithm. At the time processing address any changes into the original file to TPA, TPA will promptly intimate to the data file thus security and data integrity is secured properly. ECDSA will not allow TPA to learn any information about the data content during the auditing process. By using batch auditing TPA decrease the time of response of audit request from multiple users.

References

[1] Wang, C., Wang, Q., Ren, K., Lou, W. (2010). Privacy Preserving Public Auditing for Storage Security in Cloud Computing, IEEE INFOCOM'10, March.

[2] Cong Wang., Sherman., Chow, S.M., Qian Wang., Kui Ren., Wenjing Lou. (2013). Privacy Preserving Public Auditing for Secure Cloud Storage, IEEE , 62 (2), February .

[3] Mell, P., Grance, T. (2009). Draft NIST Working Definition of cloud computing, http://csrc.nist.gov/groups/SNS/cloud computing/index.html, June.

[4] Juels, A., Burton, J., Kaliski, S. (2007). PORs: Proof Of Retrieviability for Large Files, Proc. ACM Conf. Computer and Comm. Security(CCS'07), p.584-597, October.

[5] Qian Wang., Cong Wang., Kui Ren., Wenjing Lou. (2011). Jin Li Enabling Public Auditability And Data Dynamics For Storage

Security in Cloud Computing in IEEE transactions on parallel and distributed systems, 22 (5).

[6] Ateniese, G., Pietro, R. D., Mancini, L.V., Tsudik, G. (2008). Scalable and Efficient Provable Data Possession, Proc. Int'l Conf. Security and Privacy in Comm. Networks (SecureComm '08), p. 1-10,.

[7] Wang, C., Wang, Q., Ren, K., Lou, W. (2012). Towards Secure and Dependable Storage Services in Cloud Computing, IEEE Trans. Service Computing, 5 (2), 220-232, April - June.

[8] Shacham, H., Waters, B. (2008). Compact Proofs of Retrievability, *In*: Proc. Int'l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (Asiacrypt), 5350, p. 90-107, December

[9] Ateniese, G., Burns, R., Curtmola, R., Herring, J., Kissner, L., Peterson, Z., D.Song, Provable Data Possession at untrusted stores, in CCS'07.

[10] Dalia Attas., Omar Batrafi. (2011). Efficient Integrity Checking Technique for Securing Client Data in Cloud Computing, in IJECS – IJENS,.

[11] Shah, M. A., Baker, M., Mogul, J. C., Swaminathan, R. (2007). Auditing to Keep Online Storage Services Honest, in Proc. 11th USENIX Workshop Hot Topics in Operating Systems (HotOS '07), p 1-6.

[12] Lamba, S., Sharma, M. (2013). An Efficient Elliptic Curve Digital Signature Algorithm (ECDSA), IEEE Conference Publication, International Conference on Machine Intelligence Research and Advancement 10.1109/ICMIRA. 41.

[13] Aqeel Khalique., Kuldip Singh. (2013). Sandeep Sood Implementation of Elliptic Curve Digital Signature Algorithm, *International Journal of Computer Applications* (0975 – 8887) 2 (2), May.

[14] Wang, C., Wang, Q., Ren, K. (2009). Ensuring Data Storage security in Cloud Computing, IEEE Conference Publication, 17th International Workshop on Quality of Service (IWQoS),.

[15] Ruj, S., Stojmenovic, M., Nayak. (2012) Privacy Preserving Access Control with Authentication for Securing Data in Clouds. In Cluster, Cloud and Grid Computing (CCGrid), 12th IEEE/ACM International Symposium on (p. 556-563).

[16] Zhou, M., Mu,Y. Privacy-Preserved Access Control for Cloud Computing, International Joint Conference of IEEE TrustCom-11/IEEE ICESS-11/FCST-11, 83–90.