International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013
ISSN 2229-5518

CLOUD STORAGE SECURITY AND PRIVACY PRESERVATION

Tushar Phalke, Shweta Kumari, Karim Shaikh, Akrati Mattoo.
S.V.P.M's COE Malegaon(Bk).
Department Of Computer Engg.

***Abstract*—Cloud computing relies on sharing of resources to achieve coherence and economies of scale similar to a utility (like the electricity grid) over a network. Using the cloud storage, users store their data on the cloud without the burden of data storage and maintenance and services and high-quality applications from a shared pool of configurable computing resources. But the problem with this scenario is that the user no longer has the data in its possession. The data can be changed at any point. So this is a formidable risk. Users should be able to just use the cloud storage as if it is local, without worrying about the need to verify its integrity. In this part the Third Party Auditor comes into picture. TPA checks the integrity of outsourced data and user becomes worry-free.**

By utilizing public key based homomorphic authenticator with random masking privacy preserving public auditing can be achieved. We further extend our result to enable the TPA to perform audits for multiple users simultaneously and efficiently. Also the support for data dynamics can be achieved.

***Index Terms*—TPA, BLS, HLA, Cloud Storage, PRF.**

I. INTRODUCTION

The basic concept of cloud computing can be understood as a cluster of normal computers taken together to get a sort of super computer and this supercomputer ofcourse does a lot of things. But cloud is not just simple collection of computer resources but it also provides a management mechanism and can serve millions of users simultaneously. The basic principle of cloud computing is to distribute the computing task to many distributed computers. Our PC handles documents, store materials, send e-mails, share files and similarly does other things for us but suppose what if our PC suddenly stops working. Definitely there will be great loss of data. But in cloud computing, cloud will do all these things for us. And thus even if our PC fails the threat to data loss would be less. Why we need cloud computing? It is to use the vacant resources of computer, improve the economic efficiency through improving utilization rate, and reduce the equipment energy consumption.

A cloud service has three different features that differentiate it from traditional hosting. On demand it is sold, typically by the minute or the hour; it is flexible - 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 CSP (the consumer needs nothing but a personal computer and Internet access)[7].

On demand it enables network access to a shared pool of configurable computing resources that can be repeatly provisioned and released with minimal management efforts or service provider. The main advantages of cloud computing is time saving but it lags in providing security. Since the data placed in the cloud is accessible to all thus security of data stored on the cloud is not granted.To ensure the security one obvious way is to adopt cryptography techniques. To maintain the reputation sometimes the CSP may hide the data corruptions.Earlier works perform auditing but do not consider the privacy protection of user's data against external auditors. To avoid this problem, we introduce in this paper an effective Third Party Auditor (TPA) to audit the user's outsourced data when needed. B. BLS(Boneh, Lynn, Shacham) algorithm is used to perform Signing[9].

To achieve effective auditing and privacy preservation random masking is applied. TPA is responsible for auditing each and every user who wish to connect with cloud. This increases the auditing time and computational overhead for TPA. In order to reduce this, the technique of bilinear aggregate signature is adopted. This helps in achieving batch auditing i.e. multiple auditing task can occur simultaneously.

II. Proposed Model

For making an efficient Third Party Auditor (TPA) which provides security and privacy, some of the requirements are fundamental and have to be met accordingly: TPA neither should demand a local copy of the data nor introduce additional burden to the cloud user. Thus in this way effective auditing of the cloud data storage takes place. The process of third party auditing should preserve user's data privacy. A public key based on Homomorphic Authenticator is combined with random masking and utilized to build a public cloud data auditing system which is privacy preserving [7].

In our system, we are considering

- Client(Cloud User): The one who has significant amount of files which he wants to store on the cloud (i.e. Cloud Server).
- CSP(Cloud Service Provider): To manage Cloud Server, having considerable storage space and to provide effective services for data storage and maintenance.
- TPA(Third Party Auditor): We are introducing TPA here as trusted one and it is trusted to assess the cloud storage service reliability on behalf of the user request.

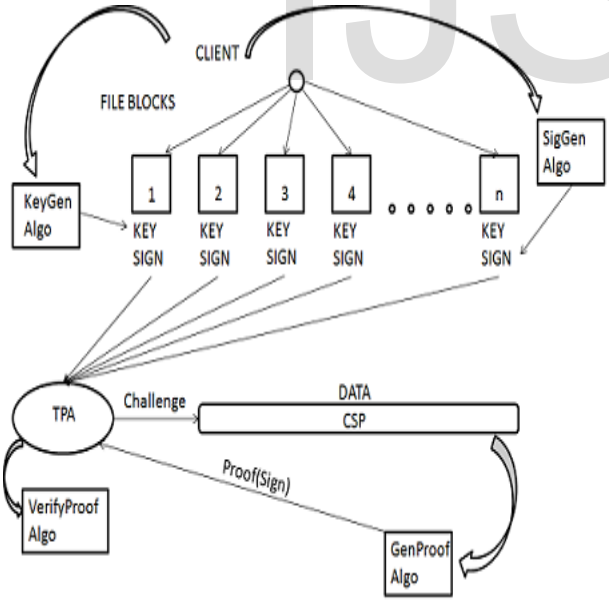

Fig. 1. Proposed Model Of Privacy Preserving In Public Auditing

For various Application Purposes users may have the need to interact with the CS in order to access and update or modify their stored data. But in our system, users will use TPA instead for ensuring storage integrity. Also by using TPA online burden will decrease and computational resources will be saved. As stated above TPA at any point will nhave no knowledge about the data. Several features like Public Auditing, Privacy Preserving, Storage Correctness and Batch Auditing are achieved.

A. Public Auditing

To allow TPA to verify the correctness of the cloud data on demand without knowing the actual data or introducing additional online burden to the cloud users, there are two phases to construct the public auditing system -Setup and Audit.The public auditing system goes through four algorithms(KeyGen algorithm, SigGen algorithm, GenProof algorithm, VerifyProof algorithm).

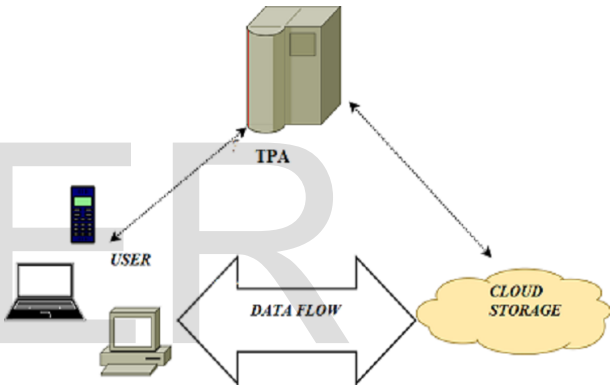

Fig. 2. Architecture Of Cloud Storage System

Setup: KeyGen algorithm is executed at Client side to achieve the public and secret parameters of the system and the data file is used by SigGen algorithm to generate the verification metadata. After this, the user's data file gets stored at the cloud. No local copy is maintained. The verification metadata is sent to TPA for later audit. Our proposed system consists of 3 phases :

Phase 1. During key generation phase, the sender chooses a random integer x $\in$ Zp and computes y=g1 $\in$ G1. The public key is y and the secret key is x.

Phase 2. Let m $\in \{0, 1\}$ be the message in the signing phase, the sender first generates h=h(m) $\in$ G1, here h () is a hash function, and then computes

σ =h^x∈ G1 . The signature of m is σ.
Phase 3. During the verification phase, the TPA first computes h=h(m) ∈ G1 and then checks whether e(h,y)=e(σ,g1).
If the verification succeeds, then the message m is authentic.

1) KeyGen Process: The user executes KeyGen to generate the public and secret arguments. Specifically, the user chooses a random signing key pair (spk, ssk), a random x ← Zp, a random element u ← G1 is chosen by the user to compute v ← gx. Here the public argument is pk = (spk, v, g, u, e(u, v)) and secret argument is sk=(x, ssk).

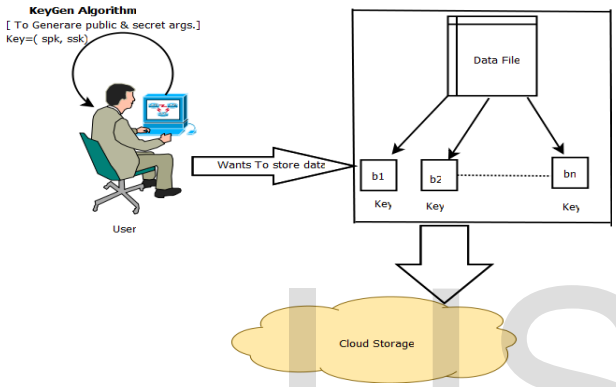

Fig. 3. Working Of KeyGen Process

Infact KeyGen process will be using a batch signature scheme which is based on BLS signature.The BLS signature scheme uses a cryptographic primitive called pairing, which can be defined as a map over two cyclic groups G1 and G2. The Boneh, Lynn, Shacham (BLS) scheme, the first short signature scheme, proposed by Boneh, Lynn, and Shacham [9] has the shortest length among signature schemes in classical cryptography. This scheme is based on Weil pairing and needs a special hash function. Weil Pairing introduced by Andre Weil is a bilinear form pairing on the points of order n of an elliptic curve taking values in nth roots of unity. It is applicable for elliptic curve cryptography.

2) SigGen Process: Suppose the data file is F = (m1, . . . ,mn), the user executes SigGen to generate authenticator i for each block mi: σi ← (H(Wi) . umi)x ∈ G1. Where Wi = name||I and name is selected by the user randomly from Zp as the identifier of file F. Let Φ =$\{\sigma i\}1\leqq i\leqq n$ be the set of authenticators.The integrity of the unique file identifier name is ensured by the last part of the SigGen. To achieve this lets compute t = name||SSigssk(name) as the file tag for F. Here SSigssk(name) is the signature of name under the secret key ssk. suppose, the TPA knows the number of blocks n. Now the user sends F to the cloud for storage along with the verification metadata .The file from local storage is deleted.

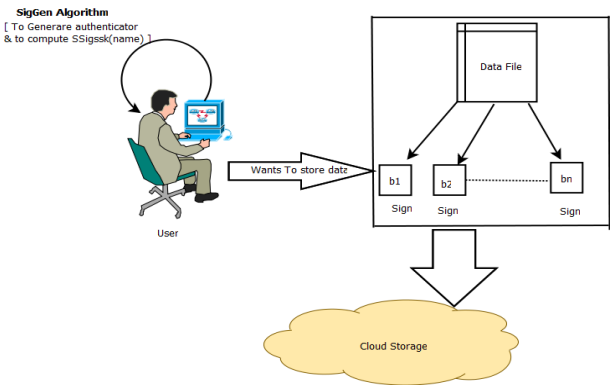

Fig. 4. Working Of SigGen Process

Audit: The TPA sends an audit message or challenge to the cloud to ensure that the data file is properly retained at the time of the audit. The cloud execute GenProof on the stored data file to generate a response message. The TPA now with the help of verification metadata, check the response by executing VerifyProof algorithm. TPA first retrieves and verifies file tag tk for each user k for later auditing. If the verification fails, TPA quits by emitting FALSE; otherwise, TPA recovers name k. Then TPA sends the audit challenge chal = $\{(i, vi)\}$i∈I to the server for auditing data files of all k users.

3) GenProof Process: When the challenge chal =$\{(i, vi)\}$i∈I is received by the CSP, it runs GenProof Algorithm which will generate a response proof in terms of the Verification Metadata for data storage correctness. A random element r ⇐ Zp is chosen by the CSP and R=e(u,v)^r∈ GT is calculated. Now the linear combination of sampled blocks in chal be denoted by μ'=Σi∈I vimi. The CSP now computes μ=r+$\mu'\gamma$ mod p where γ=h(R)∈Zp.

An aggregated authenticator $\sigma=\Pi i\in I\ \sigma^{vi}i \in G1$ is calculated by the CSP. As a response proof of storage correctness the CSP sends $\{\mu,\sigma,R\}$ to the TPA.

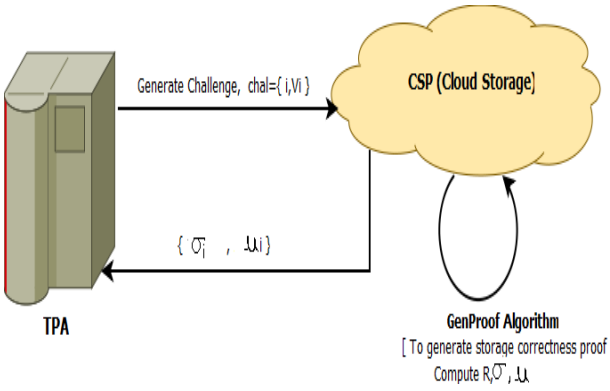

Fig. 5. Working Of GenProof Process

4) VerifyProof Process: When the response from the CSP is received by the TPA, it runs VerifyProof Algorithm to verify that response. It first computes h(R) and then checks the verification equation $R.e(\sigma^{\gamma}$,g) =?e((Π H(Wi)^{vi})$^{\gamma}$ .u$^{\mu}$,v).It thus verifies that the data is authentic i.e. free from any kind of modification.

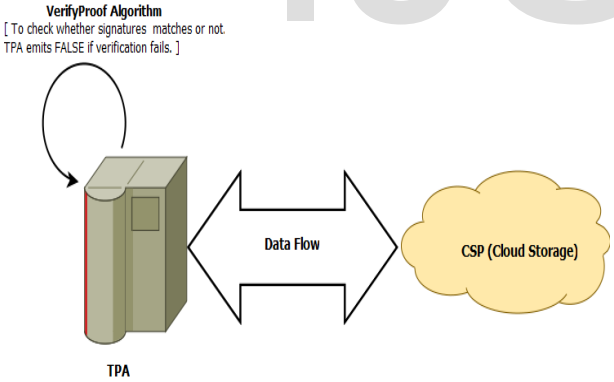

Fig. 6. Working Of SigGen Process

B. Privacy preserving

Privacy preserving can be understood as the data user wants to store on the cloud must not be revealed to any other resource. Even the TPA involved in the auditing process should not be allowed to see the actual data user wants to store.To store the data on the cloud there are two possible ways that can be chosen. The first one is a MAC-based solution which has some undesirable systematic demerits bounded usage and stateful verification in a public auditing setting.In short the mac based is not efficient enough to ensure the integrity of data.here the TPA is statefull i.e. the data is revealed to TPA.So the other way by which privacy preserving can be achieved is to uniquely integrate the homomorphic authenticator with random masking technique.

In HLA[3] the linear combination of sampled blocks in the servers 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 users 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[3]. We can also ensure the privacy preserving of the system by Theorem : Our batch auditing protocol achieves the same storage correctness and privacy preserving guarantee as in the single-user case.

C. Storage Correctness

If suppose, data on the cloud is get modified then the verification metadata generated by CSP cannot matches with original signature of TPA. Otherwise if it matches then we can say that the data stored on the cloud is correct i.e. the data is not modified. We need to prove that the cloud server can't generate valid response for the TPA without faithfully storing the data, as captured by Theorem: If the cloud server passes the Audit phase, then it must indeed possess the specified data intact as it is[1].

D. Batch Auditing

To enable TPA with secure and efficient auditing capability for multiple users simultaneously, there are K users having K files on the same cloud. They will have the same TPA. The TPA can combine their queries and save the computation time. The comparison function that compares the aggregate authenticators has a property that allows checking of multiple messages in one equation. So simply,

instead of 2K operation, K+1 operations are possible. Batch auditing protocol proposed by Wang[1] achieves the same storage correctness and privacy preserving guarantee as in the single-user case.

E. Support for Data Dynamics

In Cloud Computing, outsourced data might not only be accessed but also updated frequently by users for various application purposes [10],[11]. Hence, supporting data dynamics for privacy preserving public auditing is also of paramount importance.

III. Comparison Between Existing System and Proposed System

A. Existing System

Since cloud service providers (CSP) are separate administrative entities, data outsourcing is actually relinquishing users ultimate control over the fate of their data. As a result, the correctness of the data in the cloud is being put at risk due to the following reasons. One of the risk is although the infrastructures under the cloud are much more powerful and reliable than personal computing devices, they are still facing the broad range of both internal and external threats for data integrity

B. Demerits Of Existing System

Although outsourcing data to the cloud is economically attractive for long-term large-scale storage, it does not immediately offer any guarantee on data integrity and availability. As users no longer physically possess the storage of their data, traditional cryptographic primitives for the purpose of data security protection cannot be directly adopted. In particular, simply downloading all the data for its integrity verification is not a practical solution due to the expensiveness in I/O and transmission cost across the network. Besides, it is often insufficient to detect the data corruption only when accessing the data, as it does not give users correctness assurance for those unaccessed data and might be too late to recover the data loss or damage.The Second main drawback of all the existing system is that the TPA always demand the users data while auditing because of this the data integrity is threatened and the privacy preserving of data is not achieved.

So the system propose in this paper is efficient enough to overcome the above stated demerits. We are utilizing and uniquely combining the public key based homomorphic authenticator with random masking to achieve the privacy-preserving public cloud data auditing system, which meets all above requirements 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.

IV. Conclusion

In this system we are proposing a way for providing security to cloud storage by maintaining data integrity and privacy preserving. We are using homomorphic linear authenticator with random masking to provide the guarantee that the TPA at any point will not have any knowledge about the actual data of the user. The users data leakage is prevented. Further we propose batch auditing for saving the time by making multiple batches for better efficiency. Comparison with existing system shows the efficiency of this system.

References

[1] Cong Wang,Qian Wang,Kui Ren,Wenjing Lou(2010) *"Privacy Preserving Public Auditing for Data Storage Security in Cloud Computing"*.

[2] Patrick Honer. *"Cloud Computing security requirements and solutions: A systematic literature review"*

[3] Jachak K.B, Korde S.K, Ghorpade P.P and Gagare G.J. *"Homomorphic authentication with random masking technique ensuring privacy and security in cloud computing."*

[4] Sunil Sanka1, Chittaranjan Hota1, Muttukrishnan Rajarajan2 *"Secure Data Access in Cloud Computing"*

[5] Bhagyaraj Gowrigolla, Sathyalakshmi Sivaji, M.Roberts Masillamani *"Design and Auditing of Cloud Computing Security"*

[6] Alexa Huth and James Cebula, *"The Basics of Cloud Computing"*

[7] Balakrishnan.S, Saranya.G, Shobana.S, Karthikeyan.S *"Introducing Effective Third Party Auditing (TPA) for Data Storage Security in Cloud"*

[8] Sedat Akleylek , Bars Bulent Krlar , Omer Sever, and Zaliha Yuce. *" Short Signature Scheme From Bilinear Pairings"*

[9] D. Boneh, B. Lynn and H. Shacham., *" Short Signatures from the Weil Pairing"*, *Advances in Cryptology - ASIACRYPT01, LNCS 2248, pp. 514-532, Springer- Verlag, 2001.*

[10] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, *" Scalable and efficient provable data possession" in Proc. of SecureComm08, 2008, pp. 110*

[11] Wang, C. Wang, J. Li, K. Ren, and W. Lou, *Enabling public verifiability and data dynamics for storage security in cloud computing, in Proc. of ESORICS09, volume 5789 of LNCS. Springer-Verlag, Sep. 2009, pp. 355370.*