Privacy-Preserving Public Auditing for Regenerating-Code-Based Cloud Storage using ABE

M. H. Ranadive[1], S. S. Mujawar[2] , Varsha Patel[3]
[1, 2] Department of Computer Engineering
[1, 2] D.Y.P.C.O.E, Ambi, Pune, India

Abstract- *Cloud computing is an newly evolving computing paradigm which enables users to remotely store their data in a server and provide services on-demand with independent on location. Most of the users of cloud and cloud service providers are from different trust domains in cloud computing. Data security and privacy are becomes critical issues for remote data storage. Before cloud users have the liberty to store sensitive data to the cloud server for storage a secure user enforced data access control mechanism is provided. Attribute-based encryption is a public key encryption that enables access control over encrypted data. The protection of data which is outsourced in cloud storage against corruptions and adding fault with checking integrity of data and failure reparation becomes difficult task. A public auditing scheme is used for checking the data integrity for the cloud data with regeneration of code-based cloud storage. This scheme can completely releases the burden of data owners about storing and maintaining their data on cloud server.*

Keywords- Cloud storage, regenerating codes, public audit, privacy preserving, proxy, ABE, AES, PHR.

I. INTRODUCTION

Cloud computing is used for enabling ubiquitous, convenient, on-demand access to a shared pool of configurable computing resources. Cloud computing with storage solutions provide users and enterprises with various capabilities to store and process their data which is stored in third party data centers. Cloud storage is now gaining popularity as it offers a flexible data outsourcing service with on demand and appealing advantages such as relief of the burden for storage management, universal data access whose location is independence and avoidance of capital expenditure on hardware, software, and personal maintenances. Usually building and maintaining specialized data centers is a quite costly. So the data storage is resided with third party cloud service provider which manages this cost. In such scenario, maintaining the privacy of users data from unauthorized users is not an easy task.

The cloud service is normally faced with a broad range of internal or external adversaries, that maliciously delete or corrupt users data but on the other hand, the cloud service providers can be act dishonestly, attempting to hide data loss and pretending that the files are still correctly stored in the cloud. Another main issue is how user could actually control sharing of data which is stored on semi trusted server. Thus there is the need comes for having a mechanism of fine-grained data access control which could work efficiently with semi trusted server. A one approach is to encrypt the data before it is outsourced to the semi trusted server. To allow data owner to decide with encryption and access mechanism is the best way. Access to original data should only be given to those users having proper decryption key. Whenever the need occurs the data owner should have right to grant and revoke access privileges. This system is a secure cloud storage system which support privacy-preserving public auditing. Attribute based encryption which is a public key encryption which enables access control over encrypted data using access policies and ascribed attributes.

A secure user enforced the data access control mechanism which must be provided before cloud users have the freedom to outsource sensitive data to the cloud for storage. With the emergence of sharing confidential corporate data on cloud servers, it is imperative to adopt an efficient encryption system which has a fine-grained access control to encrypt outsourced data. Personal health record (PHR) is an emerging patient-centric model of health information exchange, which is often outsourced to be stored at a third party, such as cloud providers due to the high cost of building and main training specialized data centers. A PHR service allows a patient to create, manage, and control their personal health information from one place through the web, which has made the storage, retrieval, and sharing of the medical information more efficient. To the high value of the sensitive personal health information (PHI), the third party storage servers are mostly the targets of various malicious behaviors which may lead to exposure of the PHI. The fine-grained data access control mechanism is required to ensure patient-centric privacy control over their own PHRs that work with semi-trusted servers. To encrypt the data before outsourcing is a feasible and promising approach. The PHR owner has to decide how to do encryption of their files and to allow which set of users to get access to each file. A PHR file should only be available to those users who are given the corresponding decryption key, while remain confidential to the rest of users. Thus there have been wide privacy concerns as personal health information could be stored to those third party storage servers and not to authorized parties. To assure the patients control over access to their own PHRs is a most promising method i.e. before outsourcing to encrypt the PHRs. Public audit-ability is enabled between the two cloud storage servers to check the integrity of outsourced data.

II. LITERATURE REVIEW

Following literature is analyzed for existing methodology working and critically evaluated on some evaluation method to find shortcomings from them.

[1]" Towards secure and dependable storage services in cloud computing"

C. Wang, Q. Wang, K. Ren, and W. Lou proposes an effective and flexible distributed storage verification scheme with explicit dynamic data support to ensure the correctness and availability of users data in the cloud. It rely on erasure correcting code in the file distribution preparation to provide redundancies and guarantee the data dependability against Byzantine servers, where a storage server may fail in arbitrary ways. This construction drastically reduces the communication and storage overhead as compared to the traditional replication-based file distribution techniques. By utilizing the homomorphic token with distributed verification of erasure-coded data, this scheme achieves the storage correctness insurance as well as data error localization: whenever data corruption has been detected during the storage correctness verification, this scheme can almost guarantee the simultaneous localization of data errors, i.e., the identification of the misbehaving servers. The proposed design allows users to audit the cloud storage with very lightweight communication and computation cost. The auditing result not only ensures strong cloud storage correctness guarantee, but also simultaneously achieves fast data error localization, i.e., the identification of misbehaving server.

[2] "Privacy-Preserving Public Auditing for Regenerating-Code-Based Cloud Storage"

Jian Liu, Kun Huang, Hong Rong, Huimei Wang and Ming Xian proposes a public auditing scheme for the regenerating-code-based cloud storage system, where the data owners are privileged to delegate TPA for their data validity checking. To protect the original data privacy against the TPA, They randomize the coefficients in the beginning rather than applying the blind technique during the auditing process. Existing remote checking methods for regenerating-coded data only provide private auditing, requiring data owners to always stay on-line and handle auditing, as well as repairing, which is sometimes impractical. Thus a proxy is used who works in the absence of data owner for solving the regeneration problem of failed authenticators. Thus data owner has no need to always stay on-line. A couple of keys generate a novel public verifiable authenticator which protect original data privacy against the third party auditor and preserve the privacy in cloud storage.

[3] "Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption"

M. Li, S. Yu, K. Ren, and W. Lou proposed a patient-centric framework and a suite of mechanisms for data access control to PHRs stored in semi trusted servers. To derive fine-grained and scalable data access control for PHRs, they influence attribute-based encryption algorithm to encrypt each patient's PHR file. They divide the users in the PHR system into different security domains that greatly reduces the key management complexity for owners as well as users. A high degree of patient's privacy is assured simultaneously by exploiting multi authority ABE. Personal health record is a patient-centric framework for health information exchange, which is always outsourced to be stored at a third party cloud storage. However, there is wide privacy concerns as personal health information could be exposed to those third party cloud servers and to unauthorized parties. This scheme provide scalable and secure sharing of personal health records in cloud computing using Attribute-Based Encryption.

[4] "Enabling data integrity protection in regenerating coding based cloud storage: Theory and implementation"

H. Chen and P. Lee design and implement a practical data integrity protection scheme for a specific regenerating code, while preserving its fundamental properties of fault tolerance and repair-traffic saving. DIP scheme is designed under a mobile Byzantine adversarial framework, and enables a client to verify the integrity of random subsets of outsourced data against malicious corruptions. It works under the simple assumption of thin-cloud storage and allows different parameters to be fine-tuned for a performance security trade-off. This implement and evaluate the overhead of DIP scheme in a real cloud storage test bed under multiple parameter choices. This further analyze the security strengths of DIP scheme via mathematical models. It demonstrate that remote integrity checking can be feasibly integrated into regenerating codes in practical deployment. This evaluate the running times of different basic operations such as Upload, Check, Download, and Repair, for different parameter choices.

[5] "A digital signature scheme secure against adaptive chosen message attacks"

S. Goldwasser, S. Micali, and R. Rivest presents a digital signature scheme based on the computational difficulty of integer factorization. The scheme possesses the novel property of being robust against an adaptive chosen-message attack: an adversary who receives signatures for messages of his choice where each message may be chosen in a way that depends on the signatures of previously chosen messages cannot later forge the signature of even a single additional message. This paper shows how to construct a signature scheme with such properties based on the existence of a "claw-free" pair of permutations-a potentially weaker assumption than the intractability of integer factorization.

[6] "Distributed data possession checking for securing multiple replicas in geographically dispersed clouds"

J. He, Y. Zhang, G. Huang, Y. Shi, and J. Cao provides theories for resolving the Finding an Optimal Spanning Tree in a Complete Bidirectional Directed Graph(FOSTCBDG) problem through counting all the available paths that viruses attack in clouds network environment. Also, This help the cloud users to achieve efficient multiple replicas data possession checking by an approximate algorithm for tackling the FOSTCBDG problem, and the effectiveness is demonstrated by an experimental study. This paper, provide a novel efficient Distributed Multiple Replicas Data Possession Checking (DMRDPC) scheme to overcome the two disadvantages of centre-oriented checking. The DMRDPC scheme first find an optimal spanning tree to define the partial order of scheduling multiple replicas data possession checking. This is a very complex task, since bandwidths have geographical diversity on different links of different replicas and the bandwidths between two replicas are asymmetric, and thus it is necessary to find an optimal spanning tree with the verifier as the root in a Complete Bidirectional Directed Graph (CBDG), which connects the verifier and all the replicas. Then, according to the scheduling partial order, the data possession checking from the verifier, who checks all of its children, is started. Those replicas that have passed the verification can go on checking the data possession of their children. If some replicas fail in the checking, they can obtain one copy from its parent before they continue checking the data possession of their own children.

III. SYSTEM ARCHITECTURE

Cloud Computing is a promising next-generation IT architecture which provides elastic and unlimited resources, including storage, as services to cloud users. In Cloud Computing cloud users and cloud service providers are almost certain to be from different trust domains. It turns out that on one hand sensitive data should be encrypted before uploading to cloud servers; on the other hand, a secure user enforced data access control mechanism must be provided before cloud users have the liberty to outsource sensitive data to the cloud for storage. Similar to any untrusted storage case, generated issue's can be resolved by using a cryptographic-based data access control mechanism. Under the multi-owner settings, a novel ABE-based framework for patient-centric secure sharing of PHRs is proposed in cloud computing environments. To address the key management challenges, It conceptually divide the users in the system into two types of domains, namely public and personal domains(PSDs). In Cloud Computing, Cloud users could access the system via various low-end devices such as mobile phones, which do not have much computation power. Therefore, the proposed access control mechanism should be efficient enough in the sense that the computation load addressed on both the PHR Admin and PHR Users should be affordable to these low-end devices. Keeping these challenges in mind, a cryptographic based data access control mechanism for Cloud Computing with ABE is proposed. In this work, the issue of user revocation is addressed by applying this technique to Cloud Computing. Public auditing is an art of protecting cloud data from unauthorized user. The PHR user will store the data on the cloud server. That data will be stored in the Google drive's spreadsheets. Same time one copy of that data will be stored on second cloud i.e. salesforce cloud server. Second cloud server is used as a backup for the system. While performing the task of public auditing PHR administrator will compare the first cloud server data with second cloud server data and display the tampered data. If administrator will found that some data is tampered then he will regenerate that data and automatically stores on particular location on the cloud server.

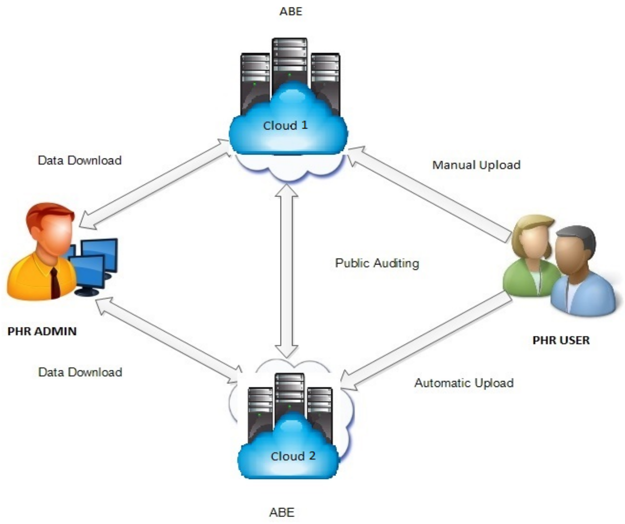


Figure 1. System Architecture

Fig 1 shows the system architecture. This architecture has four necessitates parties in a network:

1) The PHR USER

The PHR USER can manually upload the data on the cloud server.

2) The cloud server (CS)

Uploaded data is encrypted by using ABE. Cloud Server will store the encrypted data. Cloud Server provides a high-quality services utilizing a number of servers with considerable storage space and computation power.

3) The PHR ADMIN

PHR ADMIN can download the data on the cloud. Before the download, data is decrypted by using ABE. PHR ADMIN can store the downloaded data on the another one cloud server.

4) The Cloud Server(CS)

PHR ADMIN has the only access of this cloud server. PHR ADMIN can upload and download the data from this cloud server. Thus the PHR USER data automatically uploaded on the second cloud.

Public Auditing is done by the PHR ADMIN. To preserve the data integrity, PHR ADMIN will compare the data with second cloud server. Thus PHR USER has no need to stay on-line continuously.

A. Algorithm

I) AES:

```
Cipher(byte in[16], byte out[16], key-array round-key[Nr+1])
Begin
byte state[16];
state = in;
AddRoundKey(state, round-key[0]);
 for i = 1 to Nr-1 stepsize 1 do
SubBytes(state);
ShiftRows(state);
MixColumns(state);
AddRoundKey(state, round-key[i]);
end for SubBytes(state);
ShiftRows(state);
AddRoundKey(state, round-key[Nr]);
End.
```

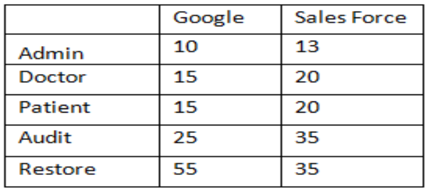
IV. EXPERIMENTAL RESULTS

Here result is discussed which will be obtained for the proposed system.

1) Cloud Communication Time Elapsed

In Fig 2 graph evaluate the cloud communication time elapsed for implementing the proposed system. following Table 1 shows the time values required for accessing the cloud server data for different modules like Admin, Doctor, Patient, Audit and Restore. The two cloud servers i.e Google and Salesforce required different time duration for access of data. Following graph shows the time required for accessing the cloud data from different cloud server. Where x-axis shows time in second and y-axis shows distinct modules.

Table 1. Time requirement in second



	Google	Sales Force
Admin	10	13
Doctor	15	20
Patient	15	20
Audit	25	35
Restore	55	35

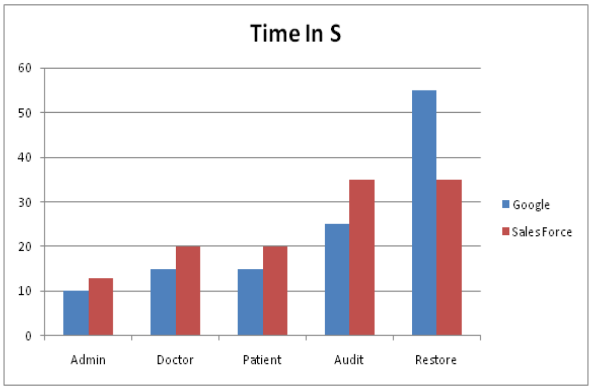
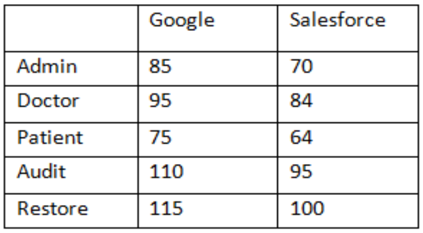


Figure 2. Time Requirement Graph

2) Memory Requirement Graph

Fig 3 graph evaluate the memory consumed for the distinct modules in the system. Following Table 2 shows the memory consumed for Admin, Doctor, Patient, Audit and Restore modules in MB for two cloud server i.e Google and Salesforce. Following graph evaluate the memory used for the various modules for implementing the system. Where x-axis shows memory in MB and y-axis shows distinct modules.

Table 2. Memory used in MB



	Google	Salesforce
Admin	85	70
Doctor	95	84
Patient	75	64
Audit	110	95
Restore	115	100

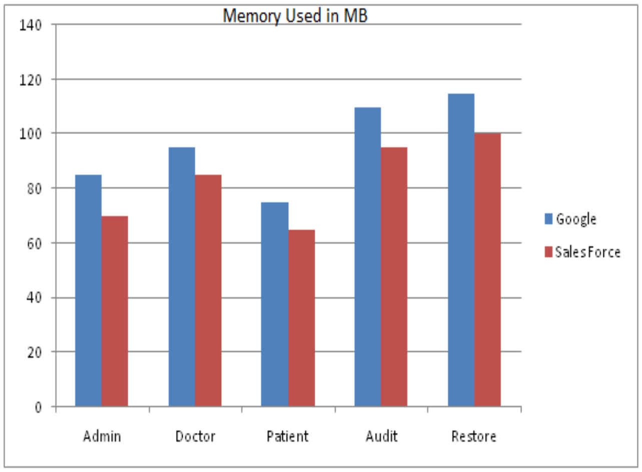


Figure 3. Memory required in MB

V. CONCLUSION

Cloud Computing is an field of plenty of challenges and of paramount importance, is still in its infancy now, and many research problems are yet to be determined. System uses attribute based encryption techniques to stores users data on the server. Every storage server has an encrypted file system who encrypts the client's data and store that data on cloud server. The system guarantees that the client's data is stored only on trusted storage servers and it cannot be accessed by intruders. Administrator can perform public auditing tasks.

VI. ACKNOWLEDGMENT

We would like to extend sincere gratitude towards our Principal Prof. A. A. Pawar(DYPCOE, Ambi, Pune), HOD (FE dept., DYPCOE, Ambi, Pune.) who have been there for constant guidance and provided support to achieve success in our endeavor.

REFERENCES

[1] Jian Liu, Kun Huang, Hong Rong, Huimei Wang, and Ming Xian. "Privacy-preserving public auditing for regenerating-code-based cloud storage". 2013.

[2] Ming Li, Shucheng Yu, Yao Zheng, Kui Ren, and Wenjing Lou. "Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption", Parallel and Distributed Systems, IEEE Transactions on, 24(1):131 ,143, 2013.

[3] Armando Fox, Rean Gri_th, Anthony Joseph, Randy Katz, Andrew Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, and Ion Stoica. "Above the clouds: A Berkeley view of cloud computing," Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, CA, USA, Tech. Rep. UCB/EECS-2009-28, 2009.

[4] Henry CH Chen and Patrick PC Lee. "Enabling data integrity protection in regenerating- coding-based cloud storage: Theory and implementation", Parallel and Distributed Systems, IEEE Transactions on, 25(2):407-416, 2014.

[5] Giuseppe Ateniese, Randal Burns, Reza Curtmola, Joseph Herring, Lea Kissner, Zachary Peterson, and Dawn Song, "Provable data possession at untrusted stores", In Proceedings of the 14th ACM conference on Computer and communications security, pages 598{609. Acm, 2007.

[6] Shafi Goldwasser, Silvio Micali, and Ronald L Rivest, "A digital signature scheme secure against adaptive chosen-message attacks", SIAM Journal on Computing, 17(2):281-308, 1988.

[7] Jing He, Yanchun Zhang, Guangyan Huang, Yong Shi, and Jie Cao,"Distributed data possession checking for securing multiple replicas in geographically-dispersed clouds", Journal of Computer and System Sciences, 78(5):1345-1358, 2012.

[8] Hovav Shacham and Brent Waters, "Compact proofs of retrievability", In Advances in Cryptology-ASIACRYPT 2008, pages 90-107. Springer, 2008.

[9] Cong Wang, Qian Wang, Kui Ren, Ning Cao, and Wenjing Lou. "Toward secure and dependable storage services in cloud computing". Services Computing, IEEE Transactions on, 5(2):220-232, 2012.

[10] R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "MR-PDP: Multiple-replica provable data possession," in Proc. 28th Int. Conf. Distrib. Comput. Syst. (ICDCS), Jun. 2008, pp. 411–420.