ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 2, Issue 5, May 2013

TRUSTWORTHINESS OF INFORMATION PROOFS STORED IN CLOUD

Arti S. Bhor[1], Smita M. Pathare[2], Khushali J. Solanki[3], Madhuri D. Dhayarkar[4]

[1,2,3,4]Dnyanganga College Of Engineering And Research

Abstract— **Cloud computing is having importance in current IT world. Cloud computing provides us with shared pool resources which we can access from anywhere without worrying about maintenance and management. It is important that data in the cloud should be correct, consistent and accessible and should have high quality. There is no guarantee that data stored in the cloud is secured and not altered by the Third Party Auditor (TPA).**

In this paper we provide scheme which gives trustworthiness of information stored in the cloud which the customer can employ to check the correctness of his data in the cloud. This proof can be agreed upon by both the cloud and the customer and can be incorporated in the Service level agreement (SLA). It also checks integrity with more accuracy. This ensures that the storage at the client side is minimal which will be beneficial for thin clients.

Keywords - Cloud computing, Trustworthiness of data, TPA, Data integrity, SLA

I. INTRODUCTION

Cloud computing has given a new dimension to the complete outsourcing arena (Software as Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS)) and they provide ever cheaper powerful processor with these computing architecture. The simplest thing that computer does is to store in available space and retrieve information whenever requested by the authenticated user. Cloud system dynamically allocates computational resources in responds customers' resource reservation requests. It helps enterprises to have a dynamically scalable abstracted computing infrastructure that is available on demand and on pay-per-use basis. Storing user data in the cloud has interesting security concerns which need to be extensively investigated to make it a reliable solution to problem avoiding local storage data. Many problems like data authentication and integrity, (i.e. how efficiently and securely the cloud storage server returns correct, complete and not modified results in the response to its clients' queries[1])outsourcing encrypted data and associated difficult problems dealing with querying over encrypted domain [2] were discussed in research literature.

From the perspective of data security Cloud Computing inevitably poses new challenges of security threats. At first, traditional cryptographic primitive for the purpose of security protection cannot be directly adopted due to the users' loss control data. Hence, we require verification of data storage in the cloud. Considering various kinds of data for each user and the demand of long term assurance of their data safety, the problem of verifying accuracy, integrity and correctness of data storage in the cloud becomes even more challenging. Secondly, Cloud Computing is not just a third party data warehouse. The stored data in cloud may be frequently viewed by the users who will perform various operations like insertion, deletion, modification, affixing, reordering, etc. Correctness of dynamically stored data is an important concern in cloud.

Data integrity is defined as the accuracy and consistency of stored data, in absence of any alteration to the data between two updates of file and record that means giving assurance to the user that his/her data is not modified by any unauthorized user. Data outsourcing is nothing but the owner (client) of the data moves its data to a third party cloud storage server which is supposed to store the data with it and provide it back to the owner whenever required. Although outsourcing of data into the cloud is economically attractive for cost and complexity of long term large scale data storage, its lacking of offering strong assurance of data integrity, availability impede its wide adoption by both enterprise and individual cloud users [3].

In this paper we deal with the problem of implementing protocol for obtaining the Proof of Retrievability (POR). This protocol tries to obtain and verify a proof that the data that is stored by a user at remote data storage in the cloud is not modified by the archive and thereby the integrity of the data is assured. Such kinds of proofs are very helpful in peer-to-peer storage systems, network file systems, long term archives, web service object stores. Such verification systems prevent the cloud storage archives from modifying the data stored at it by using frequent checks on the storage archives. Such checks must allow the data owner to efficiently, frequently, quickly and securely verify that the cloud archive is not cheating the owner where cheating means modifying some of the data.

II. RELATED WORK

Juels and Kaliski [4] proposed a scheme called Proof of Retrievability (POR). POR is for huge size of files named as sentinels. Sentinels play an important role when cloud needs to access only a small portion of the file (F) instead of accessing entire file. Sravan and saxena[5] proposed a schematic view of a Proof of Retrievability based on inserting random sentinels in the data file. Semantic view of POR is shown in Fig.1.

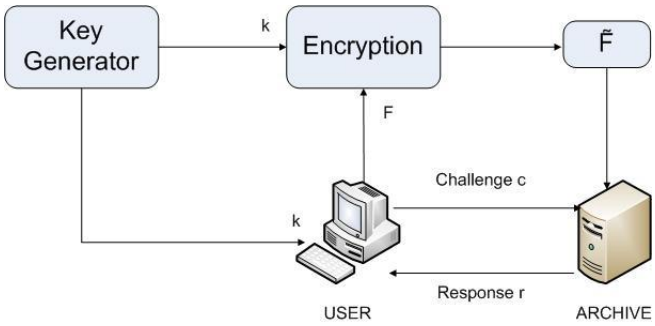

Fig.1 Schematic view of a proof of retrievability based on inserting random sentinels in the data file F [6]

The above architecture describes that the user (cloud client) stores a file (F) in the cloud server (archive). Before storing the file to the cloud, owner needs to encrypt the file in order to prevent it from the unauthorized access.

A) Data Integrity Proof In Cloud:

Cloud storage can be attractive means of outsourcing the day-to-day management of data, but ultimately the responsibility and liability for that data falls on the company that owns the data not the hosting provider. It is important to understand some of the factors like causes of data corruption, how much responsibility a cloud service provider holds, some best practices for utilizing cloud storage safely, and some best methods and standards for providing the integrity of data regardless of whether that data resides locally or in the cloud.

Integrity checking is essential in cloud storage as providing integrity is critical for any data center. Data corruption can happen at any level of storage and with any type of media. Bit rot controller failures, reduplication metadata corruption and tape failures are all examples of different media types causing corruption. Metadata corruption can be the result of any of the vulnerabilities listed above, such as bit rot, but are also susceptible to software glitches outside of hardware error rates. Unfortunately, a side effect of reduplication is that a corrupted file, block, or byte affects every associated piece of data tied to that metadata. The truth is that data corruption can happen anywhere within a storage environment. Cloud storage systems are still data centers, with hardware and software, and are still vulnerable to data corruption. One needs to look the recent highly publicized Amazon failure. Not only many companies suffered from prolonged downtime, but 0.07 percent of their customers actually lost the data. It was reported that this data loss was caused by recovering an inconsistent data snapshot of Amazon ESB volumes.

B) Secure Data Computation Outsourcing In Cloud:

Fundamental concern to move computational workloads from private resources to the cloud is the protection of confidential data that computation consumes and produces. Secure computation outsourcings services are in great need to not only protect sensitive workload information but validate the integrity of the computational result. This is, however very difficult task due to number of challenges that have to be met simultaneously. Firstly such a service has to be practically feasible in terms of computational complexity. Secondly, it has to provide sound security guarantee without restrictions of system assumptions. Thirdly, it should enable substantial computational savings at the end-user's side as compared to the amount of the efforts that otherwise has to be committed to solve the problem locally.

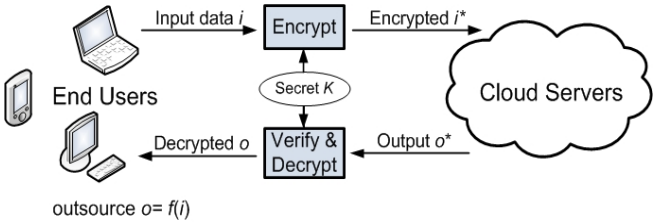

Fig.2 Secure Data Computation Outsourcing In Cloud

Our methodology is to decompose computations into public programs and private data and leverage structures of specific computations achieving desirable trade-offs security, efficiency, and practicality. We plan to organize secure outsourcing mechanisms into hierarchy, where computation can be represented at various abstraction levels, such that the aforementioned trade-offs can be flexibly explored in a systematic manner. Two critical applications to be studied are that this project includes secure outsourcing systems of linear equations [6] and secure outsourcing linear programming [7] in the cloud. These two applications are among the most widely used algorithmic and computational tools in various engineering disciplines that analyze and optimize real-world systems. The study would prepare a solid knowledge base and provide insights for further research on more advanced computation problems, such as secure outsourcing convex programming in cloud.

III. PRAPOSED WORK

The current section discusses various aspects that should be considered to achieve data integrity. Company who wishes to go for cloud storage service must be an authorized user and register themselves as a client. After registering, every authorized user will have his/her account. Fig.3 shows the register page for new user. The same account will be used by user to login.

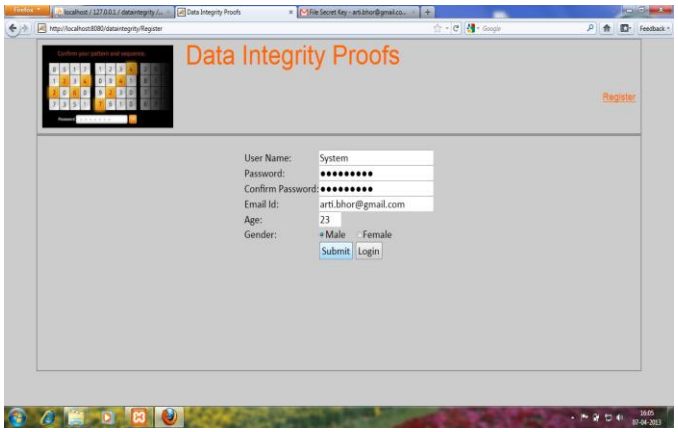

Fig.3 Register page for New User

After login the user can see the home page as shown in Fig.4 where he will have the different options related to the documents.

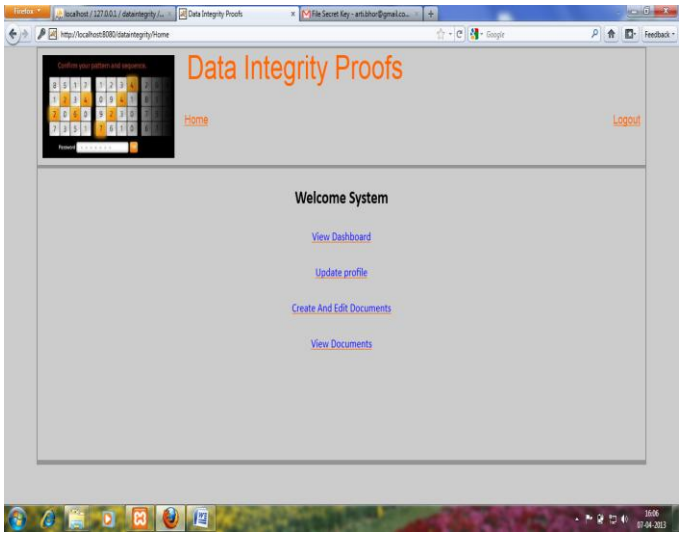

Fig.4 Data Integrity Proofs Home Page

Whenever he wishes to store any data on cloud, the system will generate the secrete key for that document. As shown in Fig.4 the secrete key will be sent to the user through the e-mail. The data in the document will be hashed and stored on to the cloud along with the secret key. The key will be independent of the data in the documents and there will be a seperate key for every document being created by an user.

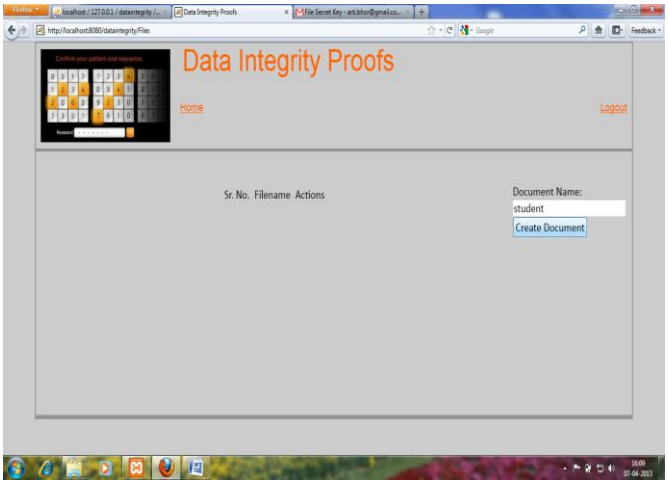

Fig. 5(a) Document Creation

The key for particular document will be required whenever the user wants to either view, edit or delete the document which increases the security level.

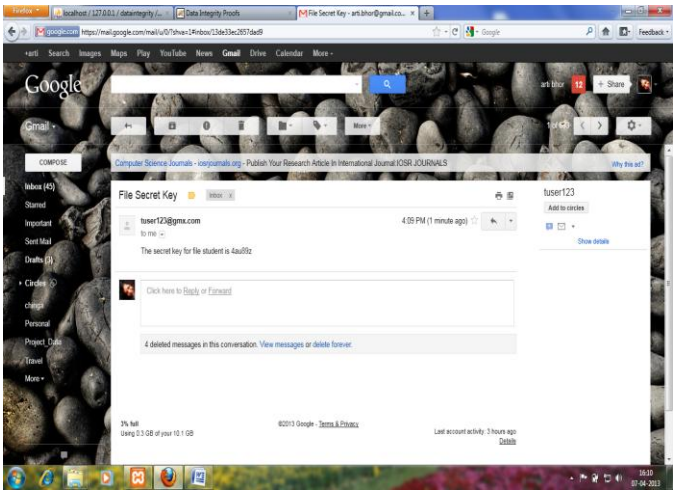

Fig. 5(b) Secret key for file through E-mail

The Fig.6(a) shows the secret key request when a user wants to edit the document he has created. Fig.6(b) shows the document window after providing the secrete key.

The proposed system ensures that an unauthorized user is not permitted to carry out any operation related to documents. For every file or document stored on cloud the system verifies whether it is secured or not. The document which is requested by user will be hashed again and that newly created hashed document will be compared with the one which was created and stored earlier. Any mismatch between the two would mean a loss of integrity.

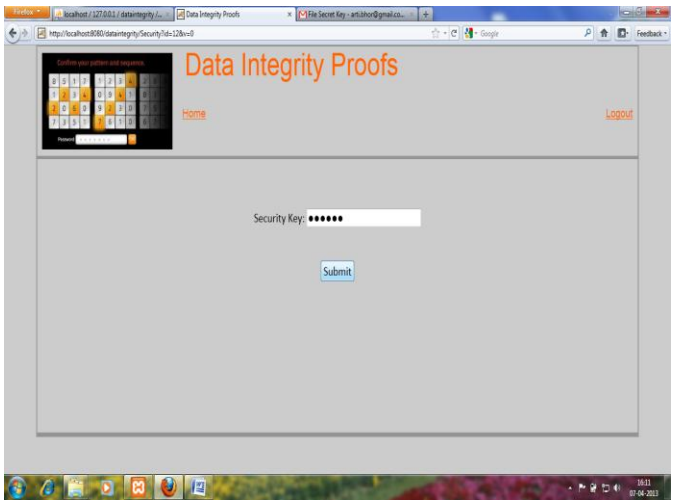

Fig.6(a) Secret key request to edit the document

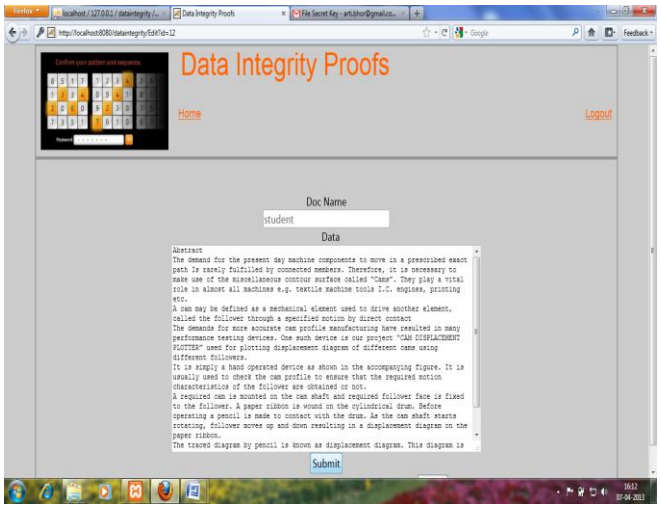

Fig. 6(b) Edited document page after providing secrete key

The user can also check the same by going through the dashboard which keeps the track of the file access time and date as shown in Fig. 7.

IV. Conclusion

In this paper we have tried to give an assurance to the clients that their data is secured in cloud server with the help of proof of data integrity hence confidentiality of users' sensitive files is maintained. This proof can be agreed upon by both the cloud and the customer and can be incorporated in the Service level agreement (SLA). Cloud Server transmits the file across the network to the user thus consumes heavy bandwidths. We present a scheme which does not involve the encryption of the whole data thus reducing the network bandwidth consumption and use of

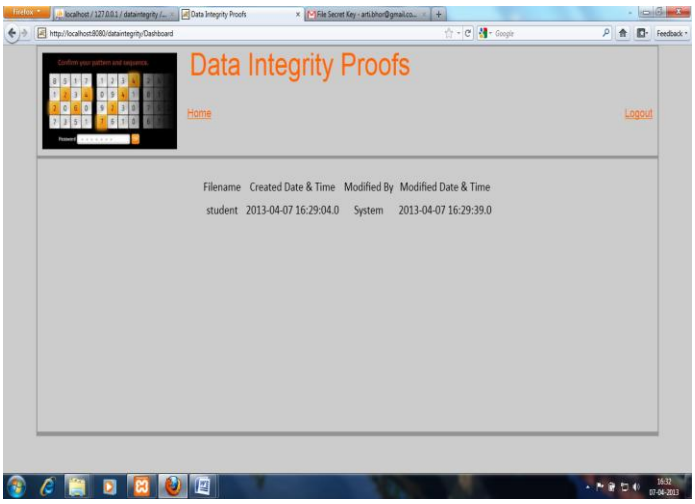

Fig. 7 Dashboard

hash function for the encryption that reduces computational overhead at client side. It also reduces the chance of losing data by hardware failures. This scheme is more advantageous to the mobile phones and PDAs which have limited CPU and battery power and communication bandwidth. It evaluates the performance of cloud storage as it consumes less computational power.

References

[1] E. Mykletun, M. Narasimha, and G. Tsudik, "Authentication and integrity in outsourced databases," Trans. Storage, vol. 2, no. 2, pp. 107–138, 2006

[2] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in SP '00: Proceedings of the 2000 IEEE Symposium on Security and Privacy. Washington, DC, USA: IEEE Computer Society, 2000, p. 44.

[3] Cong Wang, Qian Wang, Kui Ren, Wenjing Lou, "Towards Secure and Dependable Storage Services in Cloud Computing," IEEEtransactions on Services Computing, 06 May 2011

[4] A.Juels and B.S. Kaliski, Jr., "Pors: proofs of retrievability for large files," in CSS' 07:Proceedings of6 the 14th ACM conference on Computer and communications Security.New York, NY, USA:ACM, 2007, pp.584-597

[5] R. Sravan kumar and Saxena ,"Data integrity proofs in cloud storage" in IEEE 2011.

[6] Cong Wang, Kui Ren, Jia Wang, and Karthik Mahendra Raje Urs, "Harnessing the Cloud for Securely Solving Large Systems of Linear Equations," The 31st International Conference on Distributed Computing Systems (ICDCS'11), Minneapolis, MN, June 20-24, 2011.

[7] Cong Wang, Kui Ren, and Jia Wang, "Secure and Practical Outsourcing of Linear Programming in Cloud Computing", The 30th IEEE Conference on Computer Communications (INFOCOM'11), Shanghai, China, April 10-15, 2011.