RESEARCH ARTICLE OPEN ACCESS

Secure Deduplication in Hybrid Cloud Using Differential Privileges

Vinutha B T [1], Yogish H K [2]
M.Tech Student [1], PG Coordinator [2]
Department of CSE
EWIT, Bangalore
Karnataka - India

ABSTRACT

Recent years have witnessed the trend of maximum usage of cloud-based services for large scale content storage, processing, and distribution. Security and privacy are main concerns for the public cloud usage. Towards these security challenges, we propose and implement, on Open Stack Swift, a new client-side deduplication scheme for securely storing and sharing outsourced data via the public cloud. The originality of our proposal is twofold. First, it ensures better confidentiality towards unauthorized users. That is, every client computes as per data key to encrypt the data that he intends to store in the cloud. As such, the data access is managed by the data owner. Second, by integrating access rights in metadata file, an authorized user can decipher an encrypted file only with his private key.

Keywords:- Put Your Keywords Here, Keywords Are Separated By Comma.

I. INTRODUCTION

Nowadays, the explosive growth of digital contents continues to rise the demand for new storage and network capacities, along with an increasing need for more cost-effective use of storage and network bandwidth for data transfer. As such, the use of remote storage systems is gaining an expanding interest, namely the cloud storage based services, since it provides cost efficient architectures. These architectures support the transmission, storage in a multi-tenant environment, and intensive computation of outsourced data in a pay per use business model. For saving resources consumption in both network bandwidth and storage capacities, many cloud services, namely Drop box, wuala and Memopal, apply client side deduplication .This concept avoids the storage of redundant data in cloud servers and reduces network bandwidth consumption associated to transmitting the same contents several times.

Despite these significant advantages in saving resources, client data deduplication brings many security issues, considerably due to the multi-owner data possession challenges. For instance, several attacks target either the bandwidth consumption or the confidentiality and the privacy of legitimate cloud users. For example, a user may check whether another user has already uploaded a file, by trying to outsource the same file to the cloud. Recently to mitigate these concerns, many efforts have been proposed under different security models, these schemes are called Proof of Ownership systems (PoW). They allow the storage server check a user data ownership, based on a static and short value (e.g. hash value). These security protocols are designed to guarantee several requirements, namely lightweight of verification and computation efficiency.

Even though existing PoW schemes have addressed various security properties, we still need a careful consideration of potential attacks such as Data Leakage and poison attacks, that target privacy preservation and data confidentiality disclosure. This paper introduces a new cryptographic method for secure Proof of Ownership (PoW), based on the joint use of convergent and the Merkle-based Tree, for improving data security in cloud storage systems, providing dynamic sharing between users and ensuring efficient data deduplication .Our idea consists in using the Merkle-based Tree over encrypted data, in order to derive a unique identifier of outsourced data. On one hand, this identifier serves to check the availability of the same data in remote cloud server. On the other hand, it is used to ensure efficient access control in dynamic sharing scenarios.

II. LITERATURE SURVEY

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy n company strength. Once these things r satisfied, ten next steps are to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration are taken into account for developing the proposed system.

In 2014---IEEE Transactions-- Side channels in cloud services: Deduplication in cloud storage. Cloud storage services commonly use deduplication, which eliminates redundant data by storing only a single copy of each file or block.

Deduplication reduces the space and bandwidth requirements of data storage services, and is most effective when applied across multiple users, a common practice by cloud storage offerings. We study the privacy implications of cross-user deduplication. We demonstrate how deduplication can be used as a side channel which reveals information about the contents of files of other users. In a different scenario, deduplication can be used as a covert channel by which malicious software can communicate with its control center, regardless of any firewall settings at the attacked machine. Due to the high savings offered by cross-user deduplication, cloud storage providers are unlikely to stop using this technology. We therefore propose simple mechanisms that enable cross-user deduplication while greatly reducing the risk of data leakage.

In 2013---IEEE Transactions—OPS: Offline Patching Scheme for the Images Management in a Secure Cloud Environment Recent years have witnessed the development of Cloud Computing. The management of images is a big problem in virtualized environment because there are quantities of Virtual Machine images being stored in a Cloud and most of them are outdated. How to detect the outdated images and patch them efficiently? In this paper, we present a prototype called OPS- Offline Patching Scheme for the Images Management in a Secure Cloud Environment. In OPS, we can detect out the outdated image quickly by a module called Collector. Then a module called Patcher will patch the outdated images. In order to patch an image efficiently, offline patching technology is considered. For the large number of images in the Cloud, parallel scheme is also used. Our experiment results show that OPS can update numerous images efficiently.

In 2010---IEEE Transactions: Private data deduplication protocols in cloud storage:

In this project, a new notion which we call private data deduplication protocol, a deduplication technique for private data storage is introduced and formalized. Intuitively, a private data deduplication protocol allows a client who holds a private data proves to a server who holds a summary string of the data that he/she is the owner of that data without revealing further information to the server.

Our notion can be viewed as a complement of the state-of-the-art public data deduplication protocols of Halevi et al. The security of private data deduplication protocols is formalized in the simulation-based framework in the context of two-party computations. A construction of private deduplication protocols based on the standard cryptographic assumptions is then presented and analysed. We show that the proposed private data deduplication protocol is provably secure assuming that the underlying hash function is collision-resilient, the discrete logarithm is hard and the erasure coding algorithm can erasure up to α-fraction of the bits in the presence of malicious adversaries in the presence of malicious adversaries. To the best our knowledge this is the first deduplication protocol for private data storage.

III. PROPOSED SYSTEM

This Project introduces a new cryptographic method for secure Proof of Ownership (PoW), based on the joint use of convergent encryption and the Merkle-based Tree, for improving data security in cloud storage systems, providing dynamic sharing between users and ensuring efficient data deduplication.

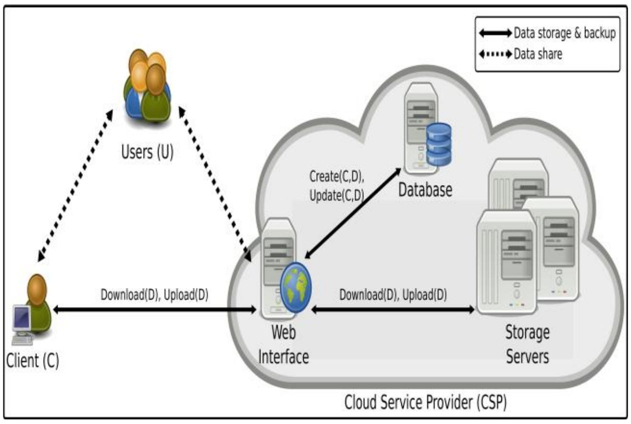
Our idea consists in using the Merkle-based Tree over encrypted data, in order to derive a unique identifier of outsourced data. On one hand, this identifier serves to check the availability of the same data in remote cloud servers. On the other hand, it is used to ensure efficient access control in dynamic sharing scenarios.

Advantages of proposed System

- A new cryptographic method for secure Proof of Ownership (PoW), based on the joint use of convergent encryption and the Merkle-based Tree improving data security in cloud storage systems.

- This identifier serves to check the availability of the same data in remote cloud servers.
- It is used to ensure efficient access control in dynamic sharing scenarios.
- Dynamic sharing between users and ensuring efficient data deduplication.
- The multi-owner data possession challenges. For instance, several attacks target either the bandwidth consumption or the confidentiality
- The privacy of legitimate cloud users.
- For example, a user may check whether another user has already uploaded a file, by trying to outsource the same file to the cloud.

Architecture



IV. FUTURE WORK

Implementation is the stage of the project when the theoretical design in turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and effective. The implementation stage involves careful planning investigation of the existing system and it’s constraints on implementation, designing of methods to achieve changeover and evaluation of changeover method.

V. CONCLUSION

The growing need for secure cloud storage services and the attractive properties of the convergent cryptography lead us to combine them, thus, defining an innovative solution to the data outsourcing security and efficiency issues. Our solution is based on a cryptographic usage of symmetric encryption used for enciphering the data file and asymmetric encryption for metadata files, due to the highest sensibility of these information towards several intrusions. In addition, thanks to the Merkle tree properties, this proposal is shown to support data deduplication, as it employs an pre-verification of data existence, in cloud servers, which is useful for saving bandwidth.

REFERENCES

[1] J. Breckling, Ed., The Analysis of Directional Time Series: Applications to Wind Speed and Direction, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[2] S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[3] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, “A novel ultrathin elevated channel low-temperature poly-Si TFT,” IEEE Electron Device Lett., vol. 20, pp. 569–571, Nov. 1999.

[4] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, “High resolution fiber distributed measurements with coherent OFDR,” in Proc. ECOC’00, 2000, paper 11.3.4, p. 109.

[5] Sorace, V. R. E. S. Reinhardt, and S. A. Vaughn, “High-speed digital-to-RF converter,” U.S. Patent 5 668 842, Sept. 16, 1997.

[6] (2002) The IEEE website. [Online]. Available: http://www.ieee.org/

[7] M. Shell. (2002) IEEEtran homepage on CTAN. [Online]. Available: http://www.ctan.org/texarchive/macros/latex/contrib/supported/IEEEtran/

[8] FLEXChip Signal Processor (MC68175/D), Motorola, 1996.

[9] “PDCA12-70 data sheet,” Opto Speed SA, Mezzovico, Switzerland.

[10] A. Karnik, “Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP,” M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.

[11] J. Padhye, V. Firoiu, and D. Towsley, “A stochastic model of TCP Reno congestion avoidance and control,” Univ. of Massachusetts, Amherst, MA, CMPSCI Tech. Rep. 99-02, 1999.

[12] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, IEEE Std. 802.11, 1997.