SECURE AND TRUSTY STORAGE SERVICES IN CLOUD COMPUTING

Saranya.V[1], Suganthi.J[2], R.G. Suresh Kumar[3]

[1,2] *Master of Technology, Department of Computer Science and Engineering,*
Rajiv Gandhi College of Engineering & Technology, (India)
[3] *Research Scholar, Vels University, Pallavaram, (India)*

ABSTRACT

Security and authentication to the data stored in the cloud is the major challenge. The paper presents major three aspects. First, it not only provides the binary status about the storage information but provides integration of the storage correctness information and data error localization (i.e. locates the misbehaving servers). Second, it supports data integrity to the dynamic operations such as update, delete and append. Third, the system is highly efficient and resilient against failures, data modifications, and even server colluding attacks. Inorder to provide the correctness and availability of the data stored in the cloud the paper provides basic tools from coding theory that is needed for file distribution across cloud servers. Then, a token which is homomorphic in nature is introduced. Subsequently, it shows the challenge response protocol for verifying the correctness and detecting the misbehaving servers. The technique for error retrieval based on erasure correcting code is also outlined. Finally, it describes how to enhance the scheme to third party auditing with only slight modification of the main design.

Keywords: Binary, Homomorphic, Integrity, Localization, Resilient

I. INTRODUCTION

In those days people look upon the sky to see the clouds moving and to seek whether there is any signs for rain (especially the farmers to accomplish the harvest). But now as the science and technology as improved and with the evolution of engineers, organizations and people look upon the cloud as a system that provides the services required by people along with storage capabilities. Several trends are opening up the era of internet based and computer technology which results in cloud computing technology. The high speed processors along with the software provides an increased pool of cloud computing services. The increase in bandwidth and reliability has resulted in transforming the information from the local machines to remote data centers. By moving the data to cloud it provides great convenience to users because they need not care about the hardware complexities in the local machines. The pioneer of Cloud Computing vendors, Amazon Simple Storage Service (S3) and Amazon Elastic Compute Cloud (EC2) are both well known examples.

This increase in internet based online services provides users with huge space for data storage but leads the users to the mercy of the cloud service providers for the availability and integrity of data. Eventhough the cloud computing services are easier and efficient than other personal digital storages the internal and external threat on the integrity of data still remains the same. Since, the users erase the copy of data from their local machines after

transforming into cloud the Cloud Service Providers (CSP) behave untrustworthy in retrieving the users data from the cloud. Inorder to increase the profit margin by reducing the cost CSPs discard the data that are least accessed by the users. This poses a greater threat on users data and both the enterprise and individual cloud users are very keen on obtaining the integrity and availability of their information from the CSPs. However, the fact that users no longer have physical possession of data in the cloud prohibits the direct adoption of traditional cryptographic primitives for the purpose of maintaining the data integrity. Hence, the cloud storage verification must be conducted without explicit knowledge of the whole data files.

The storage in cloud is not a data warehouse users also perform the primitive operations such as update, append and deletion on the data provided in the cloud. Thus, it is needed to integrate this feature along with the correctness of data provided in the cloud which is an added challenge. The deployment of Cloud Computing is powered by data centers running in a simultaneously distributed manner. It is more advantage for individual users to store their in a redundant manner to overcome the integrity threats. Thus, protocols for storage correctness assurance will be of most importance in achieving robust and secure cloud storage systems.

The description of the entire paper is as follows: After surveying the related works in Section 2, the services and technology offered by the cloud and threats that follow it are dealt in detail in Section 3. The problem definition of the proposed system is dealt in Section 4. The analysis of the proposed system is explained in Section 5. Finally, the conclusion of the paper is provided in Section 6.

II. RELATED WORKS

Juels et al. described the "Proof Of Retrievability" (POR) model that combines spot checking and error correcting code that ensures possession and retrievability of data from archived machines which in turn ensures data integrity. Shacham et al. built a model and constructed a random linear function based homomorphic authenticator which enables unlimited number of challenges and requires less communication overhead [1]. Bowers et al. proposed a framework that generalizes both Juels and Shacham models.

Later, in the subsequent works Bowers extended the POR model to distributed systems. All these models are designed for static data. The models are effective for the preprocessing steps that rest upon the file before the user outsources it. Any changes made to the file must go through the error correcting code and the shuffling process which increases the computation and communication overhead. Recently Dodis et al. gave theoretical studies on the different variants of the POR model. Ateniese et al. proposed the "Provable Data Possession" (PDP) model for ensuring the possession of the file by using a public key based homographic tags for auditing the file. The pre-computation of the tags impose a heavy overhead and it is costly.

In the subsequent works Ateniese et al. proposed a PDP scheme that uses only the symmetric key and also allowed the primitive operations on the file such as insertion, deletion and updation. This reduces the computation overhead and it is better than the previous model proposed by Ateniese et al. but, it works on only a single server. So if any crash occurs the data will be lost. Thus it leaves both the distributed servers and data availability scheme unexplored. The support of data dynamics was further studied. Wang et al. combined the homomorphic authenticator with the Merkley hash tree and provided the fully data dynamics, while Erway et al.

proposed a skip list based scheme with full data possession and dynamics [3]. The cryptography work done by Bellare et al. proposed a set of cryptographic functions such as hash, MAC and signature which enables storage integrity while performing dynamic operations on data. Curtmola et al. worked on providing multiple replicas of data across distributed systems.

Lillibridge et al. proposed a scheme in which blocks of data are dispersed across peers using an erasure code. Peers can request the data blocks from their backup peers and verify integrity using the keyed hash function located on each block. This scheme can identify the data loss but cannot guarantee that the data is unchanged . Filho et al. proposed RSA based hash to provide uncheatable data possession in peer-to-peer networks to ensure data integrity [4] [5]. Schwarz et al. proposed static file integrity across distributed servers using the erasure coding and block level file integrity techniques. In this paper we overcome all the drawbacks of the previous models by inclusion of many technical aspects.

III. CLOUD COMPUTING TECHNOLOGY

The three main services provided by the cloud are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service(SaaS). In IaaS compute, storage, networking are provided by the IaaS providers. Users use them via internet, VPN or network connection and pay for their usage. In PaaS all the hardware, software that are needed for the cloud based applications are provided by the PaaS providers and the users use them for pay and it can be utilized throughout the lifecycle. In SaaS the software runs on the computers owned by the SaaS providers and installed and managed on computers owned by users. Apart from these services there are various other services such as Communication as a Service, Utility as a Service, Security as a Service, etc. The threats on cloud computing services include loss of data, traffic hijacking, insecure interfaces and APIs, Denial of Service (DoS), malicious insiders, cloud abuse, insufficient diligence, technology vulnerabilities.

3.1 Understanding the Types of Clouds

The different types of cloud includes: Public cloud, Private cloud, On-premise private cloud, Externally hosted private cloud and Hybrid cloud. Public cloud is provided to all individual users at a “pay-for-use” rate. This enables the users to share the services of the cloud and store their information in the cloud. Private cloud is provided to individual enterprise.

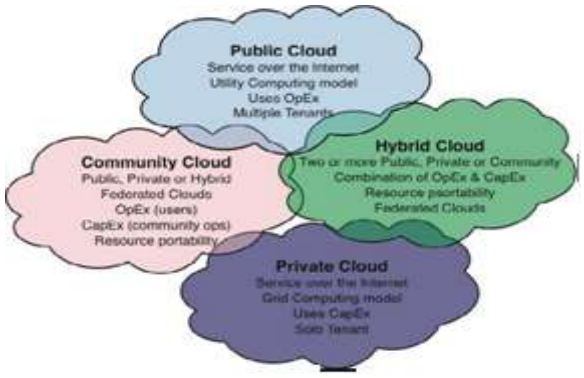

Fig.1. Types of Clouds

The aim is to share and store the enterprise information securely. On-premise private cloud also known as internal cloud is hosted in one's own data center. This provides more security of information but it is not scalable. Externally hosted private cloud is hosted externally with the cloud providers. The cloud providers provide an exclusive cloud with full security. Hybrid cloud is a combination of the private and public cloud. With the hybrid cloud the cloud providers utilize the third party in a partial or full manner thus increasing the flexibility of computing.

3.2 Challenges of Cloud Computing

The challenges of the cloud include Data Protection, Data Recovery and Availability, Regulatory and Compliance Restrictions, Management Capabilities. Data Protection refers to the security which the enterprise requires from the cloud providers. In the existing model the enterprise build their own firewalls at their data centers to protects their confidential information. In the cloud model the cloud providers are responsible for maintaining the privacy for the enterprise and the enterprise has to rely upon the cloud providers. Data recovery and availability refers to the refers to recovery during disaster and availability of data, replication of data, runtime governance and performance management.

Even though there may be different cloud providers management capabilities are still in its infancy. Features like "auto-scaling" is a crucial requirement. Providers need to work upon this and load balancing factor. The government of European countries requires the personal information of the people to store in a exclusive data center within the country and provide information when required at high privacy. This is a big challenge to the cloud providers.

IV. PROBLEM DEFINITION

4.1 System Model

Three entities are play a major role in the cloud service storage architecture: Users, Cloud servers and Third Party Auditor(TPA). Users are those who have their storage in the cloud and rely upon the cloud providers for security assurance and data integrity. Cloud servers are maintained by the Cloud Service Providers and they are responsible in providing the availability and data integrity of the information of the users. TPA are trusted third party who on behalf of the cloud users expose the risk of the cloud storage. They have capabilities that the users do not have.

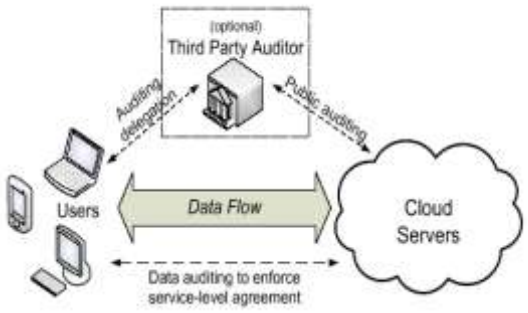

Fig.1. Cloud Storage Service Architecture

Cloud users store their data in a redundant manner because there is huge availability of data. This can be achieved by erasure correcting code to impose fault tolerance as the user data grows in size and importance. The point-to-point communication between the CSP and the users are reliable and secure. Users perform block level operations and our model provides file oriented features rather than non-file oriented features such as social networking.

4.2 Adversary Model

From the user's point of view adversary model refers to the integrity of user's data from threats. There are two types of attacks the CSPs face in maintaining the integrity of user's data: Internal and External attack. Internal attack refers to the malicious threat within the system. Inorder to overcome it the CSPs may move the less accessed data to a lower tier or even may hide it due to management errors. The external attack is due to certain parties outside the CSPs boundary who may attack at economical interest. The CSPs do not have control over the these threats.

Adversary is responsible for polluting the data and introducing fraudulent data instead of the original data and it becomes impossible for the users to access it. This also refers to external threat. Our model provides facilities that prevents fraudulent attack by the attackers and provides integrity of users data.

V. ANALYSIS OF THE SYSTEM

5.1 File Distribution Preparation

Erasure correcting code allows files to be distributed across distributed servers. We use a technique to disperse the data file F redundantly across a set of n = m+ k distributed servers. An (m, k) Reed-Solomon erasure-correcting code is used to create k redundancy parity vectors from m data vectors in such a way that the original m data vectors can be reconstructed from any m out of the m+k data and parity vectors. By placing each of the m+k vectors on a different server, the original data file survives the failure of any k of the m+k servers without data loss and overhead.

5.2 Challenge Token Pre-Computation

The assurance of data storage correctness and data error localization is achieved simultaneously by pre-computed verification tokens. Before file distribution the user pre-computes a certain number of short verification tokens on individual vector G(j) ($j \in \{1, \ldots, n\}$), each token covering a random subset of data blocks. When the user make sure the storage correctness by challenging the cloud servers with a set of randomly generated block indices. Upon receiving challenge, the cloud server computes a short "signature" over the specified blocks and returns them to the user.

5.3 Correctness Verification and Error Localization

Error localization is important for eliminating the erroneous servers. Other systems only provide binary results. But in our model we use the challenge response protocol to eliminate the erroneous servers and thus protect the users information and provide data integrity and availability.

Fig.3. System Architecture

VI. CONCLUSION

In our paper we provide the protocol extension for privacy-preserving and also discuss the cloud storage services. We have included correctness analysis of proposed storage verification design. We have presented detailed discussion on the strength of our bounded usage for protocol verifications and its comparison with state-of-the-art.

In future we revise the challenge response protocol and privacy preserving protocol to achieve better communication and space overhead.

REFERENCES

[1] Security Issues in Cloud Computing: The Potentials of Homomorphic Encryption Aderemi A. Atayero, Oluwaseyi Feyisetan.

[2] Sarvesh Kumar, Nilesh Kumar Dubey Cloud computing (A Survey on Cloud Computing Security Issues and Attacks in Private Cloud). Issue 3, volume 1(January 2013) ISSN 2249-6149.

[3] M. Jensen1, J. Schwenk2, N. Gruschka3, and L. Lo Iacono4,"On technical security issues in cloud computing," in Proceedings of the IEEE International Conference on Cloud Computing (CLOU-II), 2009.

[4] B. R. Kandukuri, R. Paturi V, A. Rakshit, "Cloud Security Issues", In Proceedings of IEEE International Conference on Services Computing, pp. 517-520, 2009.

[5] S. Dhar, "From Outsourcing to Cloud Computing: Evolution of It Services," Management Research Review, Vol. 35, No. 8, 2012, pp.664-675.

[6] A. Kundu, C. D. Banerjee, P. Saha, "Introducing New Services in Cloud Computing Environment", International Journal of Digital Content Technology and its Applications, AICIT, Vol. 4, No. 5, pp. 143-152, 2010.

[7] Journal of Cloud Computing: Advances, Systems and Applications 2012, 1:17 doi:10.1186/2192-113X-1-17