Volume 3, Issue 6, June 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Data Storage and Data Integrity in Multi-Cloud Storage

Priyanka V.Mogre
Department of CSE, RTMNU
Nagpur, India

Prof.Girish Agarwal
Department of CSE, RTMNU
Nagpur, India

Prof.Pragati Patil
Department of CSE, RTMNU
Nagpur, India

Abstract— Cloud computing is used to store data from various resources by the user. It is difficult for the user to store entire data within the system; therefore clouds are formed to store the user data. User can store as much large amount of the data as user wants. This data stored in the cloud must be integrated, the integrity of the data is thus has to be checked and maintain with the help of Trusted third party. Only trusted third party has the authority to check and to maintain the integrity f the data. The main approach of this paper is to check the integrity of the data stored and to maintain the security by using cryptography method.

Keywords— Multi-cloud storage, data integrity, cloud privacy, TTP, Cryptography, proofs of retrievability.

I. INTRODUCTION

In Cloud computing storing and sharing of data is been done via trusted third party. The trusted third party have the authorized key which help to access data between the clouds these helps & maintains privacy via TTP .Data from the various users from various resources are stored in multi-cloud. Accordingly users can login into the cloud and feed the data, user have to buy the clouds as per the requirements on the amount of data to be stored. The only thing was the cloud computing lacks regarding the issues of data integrity, data privacy, and data accessed by unauthorized members.

Data integrity: Data integrity refers to maintaining consistency of the data all over the cycle. Data integrity contains protocols for data retention specifying the length of data that can be retained. To achieve data integrity it specifies what can be done with data values when its validity expires. These protocols are consistently and routinely applied to all data entering the system and any enforcement of relaxation will cause error in the data. Strict enforcement of data integrity rules causes the error rates to be lower, resulting in time saved troubleshooting and tracing erroneous data and the errors it causes algorithms.

Three types of integrity constraints are an inherent part of the relational data model: entity integrity, referential integrity and domain integrity

- Entity integrity: It concerns with concept of Primary keys.
- Referential integrity: It concerns the concept f foreign keys.
- Domain integrity: It specifies that all the columns in database must be declared upon a defined domain.

Cloud storage: Cloud storage is an industry term for managed data storage through hosted network service. The most basic form of cloud storage allows users to upload individual files or folders from their personal computers to a central Internet server. This allows users to make backup copies of files in case their originals are lost. Users can also download their files from the cloud to other devices, and sometimes also enable remote access to the files for other people to share. Modern day cloud storage is base on highly visualized infrastructure and as the same characteristics as cloud computing in terms of elasticity, scalability and multi-tenancy, and is available on Amazon EC2 and ViON capacity service.

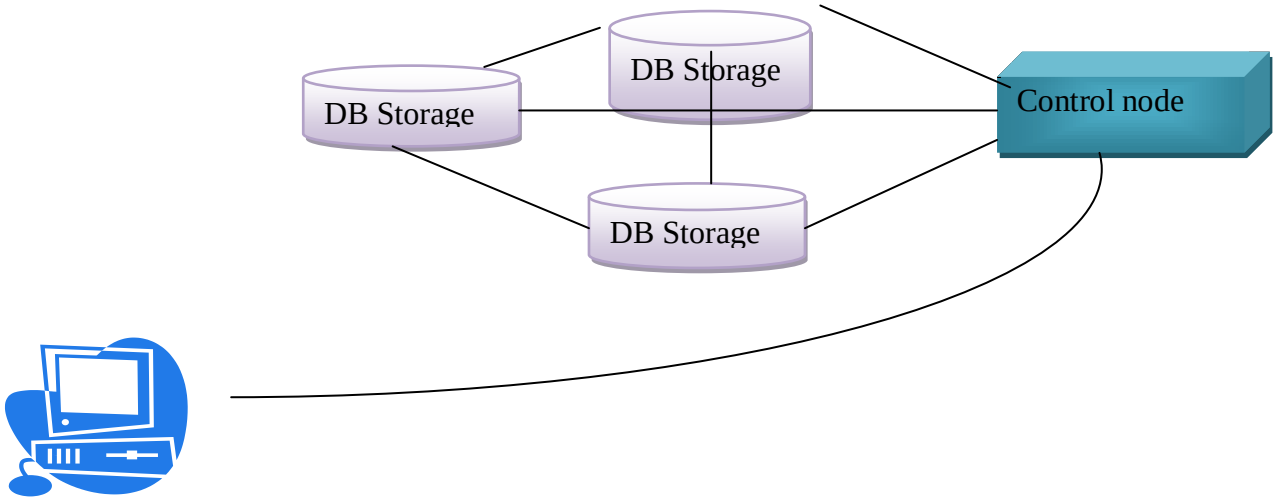

Fig.a: **Architecture of a cloud storage system including control severs and control storage servers.**

The architecture of cloud storage as shown in the above fig describes how data are stored in the database storage and the control node helps the all the database storage to store data from the system or any other resources. The benefits of cloud storage are that the maintenance task such as backup and additional storage devices are been reduces. User does not need to install any physical device at their data-centre. User only have to pay for the storage that been required to them. One way to protect against threat t web application and data is to deploy a web application Firewall as a software solution. When deployed correctly a web application firewall protects web application and data from known threats including Path Traversal, Remote Command Execution, and Compromised Servers. But the firewall must consume CPU cycle reading for each packet. This process requires more processing power which became bottleneck for the network. This means application Firewalls are less suited for real-time application but cloud storage is well suited for accessing large and huge files an complicated real time application.

II. LITERATURE REVIEW

In Cloud computing the issue of data security is still carried on, many researchers are still working with many different solutions. According to Saranya Eswaran and Dr.Sunitha Abburu Third Party Auditors can understand the Threats and they know best practices to identify the threats. Also they have the resources to check for process adherence and service quality. The TPA will be able to verify over any threats in online storage services that are represented in the cloud server.Thus, the user who owns the data can rely on the TPA to verify the data in the cloud without involving with the procedure. [1] Author: Jia Xu and Ee-Chien Chang explore Proofs of Retrievability (POR) is cryptographic method for remotely auditing the integrity of files stored in the cloud, without keeping a copy of the original files in local storage. [4] Author: Yan Zhu, Hongxin Hu, Gail-Joon Ahn, focused on Cooperative Provable Data Possession for Integrity Verification in Multi-Cloud Storage. Provable data possession (PDP) is a technique for ensuring the integrity of data in storage outsourcing. In this paper, the construction of an efficient PDP scheme for distributed cloud storage to support the scalability of service and data is addressed. In which the existence of multiple cloud service providers to cooperatively store and maintain the clients' data. This paper present a cooperative PDP (CPDP) scheme based on homomorphic verifiable response and hash index hierarchy.It also prove the security based on multi-prover zero-knowledge proof system, which can satisfy completeness,knowledge soundness, and zero-knowledge properties [2]. ZHU Yan,WANG HuaiXi, HU ZeXing, AHN Gail-Joon & HU HongXin ellaborated the Zero-knowledge proofs of retrievability .Proof of retrievability (POR) is a technique for ensuring the integrity of data in outsourced storage services. In this paper, the address to the construction of POR protocol on the standard model of interactive proof systems. This paper also proposes the first interactive POR scheme to prevent the fraudulence of prover and the leakage of verified data. It also gives full proofs of soundness and zero-knowledge properties by constructing a polynomial time rewindable knowledge extractor under the computational Diffie-Hellman assumption. In particular, the verification process of this scheme requires a low, constant amount of overhead, which minimizes communication complexity [5]. S. P. Jaikar & M. V. Nimbalkar explains Securing Cloud Data Storage. According to them Innovations are necessary to ride the inevitable tide of change. Most of enterprises are striving to reduce their computing cost through the means of virtualization. This demand of reducing the computing cost has led to the innovation of Cloud Computing [6]. According to the authors Nouha Oualha Jean Leneutre, Yves Roudier Verifying remote data integrity in peer-to-peer data storage the protocols have been proposed as a primitive for ensuring the long-term integrity and availability of data stored at remote untrusted hosts [7].

III. TECHNIQUES FOR DATA INTEGRITY

User must be able to get the data exactly stored in the cloud via trusted third party.

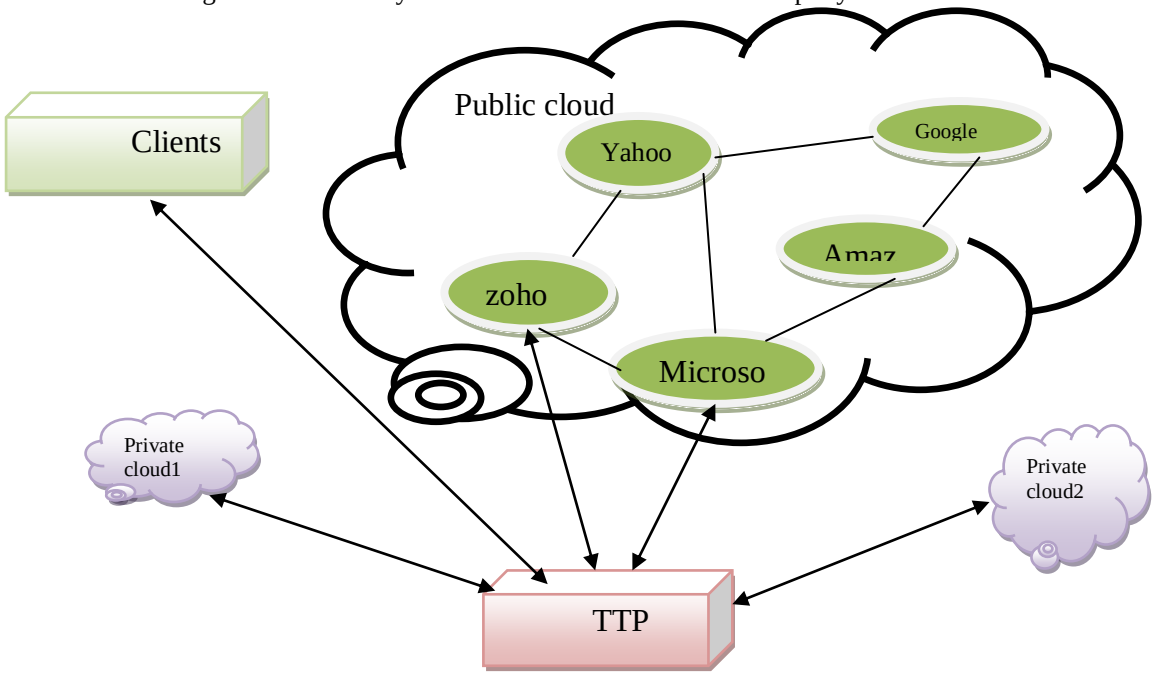

Fig.b Architecture for Data integrity

In the above architecture multiple service providers are been considered for storing user data. Moreover, a cooperative PDP is used to verify the integrity and availability of their stored data in all CSPs. The verification procedure is described as follows: Firstly, a client (data owner) uses the secret key to pre-process a file which consists of a collection of n blocks, generates a set of public verification information that is stored in TTP, transmits the file and some verification tags to CSPs, and may delete its local copy; Then, by using a verification protocol, the clients can issue a challenge for one CSP to check the integrity and availability of outsourced data with respect to public information stored in TTP [2]. For the verifying the integrity of the data stored in the cloud proof of Retrievability and cooperative proveable data possession framework is to be defined. Proof of Retrievability (POR) is designed in such a way that all the large data or files can be handled. It may screen as a cryptographic proof of knowledge technique which keeps the user data consistent.

Definition: [3] An interactive proof of retrievability scheme S is a collection of two
Algorithms and an interactive proof system, S = (K, T, P):
KeyGen(1κ): It takes a security parameter κ as input, and returns a secret key sk or a public-secret keypair (pk, sk);
TagGen(sk, F): It takes as inputs the secret key sk and a file F, and returns the triples (ζ, ψ, σ), where ζ denotes the secret used to generate the verification tags, ψ is the set of public verification parameters u and index information χ, i.e., ψ = (u, χ); σ denotes the set of verification tags;
Proof(P, V): It is a protocol of proof of retrievability between a prover (P) and a verifier (V). At the end of the protocol run, V returns {0|1}, where 1 means the file is correctly stored on the server. It includes two cases:
•P(F, σ), V (sk, ζ) is a private proof, where P takes as input a file F and a set of tags σ, and V takes as input a secret key sk and a secret of tags ζ;
•P(F, σ), V (pk, ψ) is a public proof, where P takes as input a file F and a set of tags σ, and a public key pk and a set of public parameters ψ are the common input between P and V , where P(x) denotes the subject P holds the secret x and P, V (x) denotes both parties P and V share a common data x in a protocol. This is a more generalized model than existing POR models. Since the verification process can be considered as an interactive protocol, this definition is not limited to the specific steps of verification, including scale, sequence, and the number of moves in protocol, so it can provide greater convenience for the construction of protocol. Further, this paper will only consider the construction of public proof protocol. Coperative provable data possession (**CPDP**): In Proveable data Possession (PDP) the client preprocesses the data and then sends it to an untrusted server for storage, while keeping a small amount of meta-data. The client later asks the server to prove that the stored data has not been tampered with or deleted, without downloading the actual data. Whereas in CPDP a security is maintained by exchange of security keys whenever the communication takes place between clouds and the user. Here an organizer in introduced which is one of the CSP that directly contact with the verifier. where the action of organizer is to initiate and organize the verification process. Often, the organizer is an independent server; the only advantage of this new multi-prover proof system is that it does not make any difference for the clients between multi-prover verification process and single prover verification process in the way of collaboration.

IV. CONCLUSIONS

In these paper we briefed, how the data are stored in multi-clouds and how the security of the data is maintained along with integrity of the data .Using various framework such as zero-knowledge proof of retrievability and CPDP the verification and security of the data is maintained within the cloud. The interaction between the authorized users and the CSP is considered in our scheme. As a future work ECC based algorithm may be a better choice but still it is challenging task because the exciting cryptographic schemes/techniques have too much restriction on the security.

ACKNOWLEDGMENT

We would like to thanks all the authors whom we referred for giving their suggestions and making the material available for us to refer.

REFERENCES

[1] Kavitha Murugesan, Shilpa Sudheendran "Ensuring User Security and Data Integrity in MultiCloud" nternational Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-2, May 2013.
[2] Yan Zhu, Hongxin Hu, Gail-Joon Ahn, Senior Member, IEEE, Mengyang Yu "Cooperative Provable Data ossession for Integrity Verification in Multi-Cloud Storage" IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS
[3] ZHU Yan , WANG HuaiXi , HU ZeXing , AHN Gail-Joon & HU HongXin "Zero-knowledge proofs of retrievability "Science China Press and Springer-Verlag Berlin Heidelberg 2011 .
[4] Jia Xu and Ee-Chien Chang National University of Singapore Department of Computer Science "Towards Effcient Proofs of Retrievability in Cloud Storage"
[5] ZHU Yan,WANG HuaiXi, HU ZeXing, AHN Gail-Joon & HU HongXin August 2011, Volume 54, Issue 8, pp 1608-1617 "Zero-knowledge proof of retrievability
[6] S. P. Jaikar & M. V. Nimbalkar IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-0661 Volume 1, Issue 6 (July-Aug. 2012), PP 43-49 "Securing Cloud Data Storage"
[7] Oualha, N., Onen, M., Roudier, Y.: A Security Protocol for Self-Organizing Data Storage.Tech. Rep. EURECOM+2399, Institut Eurecom, France (2008)

[8] Oualha, N., Roudier, Y.: A game theoretic model of a protocol for data possession verification. In: IEEE International Workshop on Trust, Security, and Privacy for Ubiquitous Computing (TSPUC'07). Helsinki, Finland (2007)

[9] A. Juels and B.S. Kaliski, Jr., "Pors: proofs of retrievability for large files," in CCS '07: Proceedings of the 14th CM conference on Computer and communications security.

[10] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in CCS '07.

[11] Dalia Attas and Omar Batrafi "Efficient integrity checking technique for securing client data in cloud computing" in IJECS-IJENS, 2011.

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

[13] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, "Scalable and efficient provable data possession," in *Proceedings of the 4th international conference on Security and privacy in communication netowrks, SecureComm*, 2008, pp. 1–10

[14] Priyanka V. Mogre, Girish Agarwal, Pragati Patil "Data Security and its techniques in Cloud Storage – A Review" Author Name et. al. / International Journal of Engineering Research and Technology Vol. 1 (02), 2012, ISSN 2278 - 181

[15] Anup Mathew " Survey Paper on Security & Privacy Issues in Cloud Storage Systems", EECE 571B, TERM SURVEY PAPER, APRIL 2012.

