ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

Scalable and Efficient Provable Data Possession

Hadassa Katta[1] Vivek Kolla[2] P Raja Rao[3]

Department of Computer Science, Department of Computer Science

M Tech Student, Dept., of CSE, QIS College of Engg., & Technology, Ongole, Prakasamdt, AP[1]

Assistant Professor, Dept., of CSE, QIS College of Engg., & Technology, Ongole, Prakasamdt, AP[2]

Project Lead, Polaris Financial Technology Ltd, Hyderabad, Ranga Reddy Dt, AP[3]

Abstract- Cloud storage has become an attractive and cost effective alternative for enterprises to outsource their valuable business data. However, there are security concerns pertaining to the integrity of data as the cloud server is treated as "untrusted". To overcome this problem many security schemes came into existence. Recently Zhu et al. presented a technique known as Provable Data Possession (PDP) for data integrity in cloud with distributed storage mechanisms. They considered multiple cloud service providers to store data in cooperative fashion. Their solution makes use of homomorphic verifiable response indeed and multi-prover zero-knowledge system for ensuring data integrity. In this paper we practically implement the PDP scheme proposed by Zhu et al. and build a prototype application to demonstrate the proof of concept. The empirical results reveal that the PDP scheme is very effective and can be used in real time multi-cloud environments.

Index Terms –Cloud computing, outsourcing, multiple cloud, and cooperative data possession

I. INTRODUCTION

Cloud computing has become popular in recent years. There are many cloud providers such as Amazon, IBM, Microsoft, and Oracle and so on. These cloud service providers help users to outsource their data to cloud. The cloud environment is built based on standards for interoperability. Thus multi-cloud environments in distributed architecture came into existence. They are also known as hybrid clouds. Virtual Infrastructure Management (VIM) is one the features of these clouds. Amazon EC2 web services are an example for such distributed environment. There are many technologies or tools for multi-cloud. They include Ovirt, vSphere, VMware, and Platform VM Orchestrator. Distributed cloud storage platforms can be built using these tools. Thus is possible for the cloud to support data outsourcing facilities to users. However, there are security concerns regarding this concept as they cloud servers are treated by the users as untrusted. The valuable business data when stored in cloud server, without a local copy, there is security concern as the cloud storage providers do not take or data security generally or they have limitations in providing security. Therefore the data stored in cloud may be vulnerable to attacks thus causing irreversible losses to clients. This is the problem to be addressed. Many schemes came into existence to address this problem. Provable Data Possession (PDP) is one such scheme proposed in [1]. This scheme ensures that the data integrity is not lost. However, this scheme needs the users to download data for verification which causes security problem again. Therefore it is essential to have a scheme where data downloading is not required for verification. Towards this end PDP scheme such as Scalable PDP [2] and Dynamic PDP [3] came into existence. These schemes focused on single cloud storage providers.

There are schemes like SPDP [4], DPDP-II [4] and DPDP-I and Merkle Hash Tree (MHT) make use of authenticated skip list in order to verify the adjacent blocks for integrity. These schemes do not work in multi-cloud environments as they can't construct MHT for such environment. The other schemes such as CPOR-I and CPOR-II [5] and PDP [4] make use of homomorphic verification tags where downloading data for verification is not required. To overcome the drawbacks of all existing methods, it is essential to build a cooperative PDP scheme that works in distributed multi-cloud environment. Data LeakageAttack and Tag Forgery Attack are the attacks to be prevented with new PDP scheme. Security models for existing PDP schemes were proposed in [4], [6], and [5]. However, they can't cover all security problems. To overcome these problems recently Zhu et al. [7] presented a novel PDP scheme which ismeant for data integrity verification in multi-cloud environments besides preventing various attacks.

In this paper we implement the PDP scheme proposed in [1]. We build a prototype application to demonstrate the proof of concept. The experimentalresults revealed that the scheme is effective and useful. The remainder of this paper is organized as follows. Section II reviews relevant literature. Section III provides verification architecture for data integrity. Section IV presents cooperative PDP scheme.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

Section V presents experimental results while section VI concludes the paper.

II. PROIR WORKS

Two approaches are found in literature to ensure integrity and availability of outsourced data. They are known as Proofs of Irretrievability (POR) [1] and Provable Data Possession (PDP) [4]. PDP uses RSA-based scheme for security. A publicly available version of PDP is also availabhle which enables any one to verify the data. It also causes problems when data owners are separated from data users. These schemes are proved to be insecure against replay attacks. Moreover, they are not suitable for multi-cloud storage environment. Other PDP schemes such as dynamic PDP and scalable PDP [2] where servers can deceive owners as the challenges lack randomness. Data leakage is another problem with these schemes besides reflecting suitability for multi-cloud environment. With respect to POR schemes, they depend on pre-processing steps conducted by client before data is sent to cloud. Then this scheme does not properly support data dynamics. An improve version of POR known as Compact POR was introduced in [5] which makes use of homomorphic property for security. However, it also could not prevent data leakage attacks. A dynamic scheme was presented in [6] whichintegrate MHT and CompactPOR into DPDP. Later many POR schemes came into existence [8], [9]. To overcome these problems of existing schemes recently Zhu et al. [7] presented a novel PDP scheme which is meant for data integrity verification in multi-cloud environments besides preventing various attacks.

III. VERIFICATION FRAMEWORK FOR MULTICLOUD

The proposed PDP scheme is implemented in multi-cloud environment. The environment includes combination of public and private clouds. Figure 1 shows the data flow for verification of integrity of data outsourced to the multi-cloud storage.

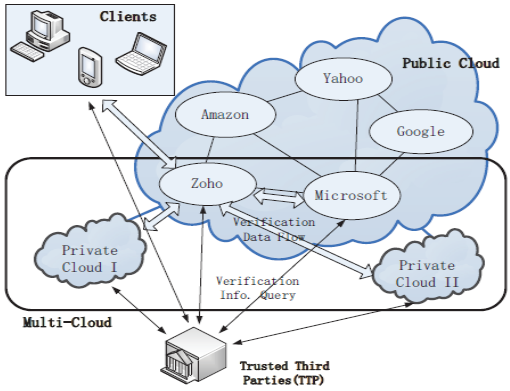

Fig. 1 –Data Integrity Verification in Multi-Cloud (excerpt from [7])

As can be seen in figure 1, it is evident that there are three parties involved. They are cloud service providers in cooperative environment, clients and trusted third parties. The trusted third party makes query for data integrityverification. The verification data flows between the cloudsand finally the TTP is able to verify the integrity of data. Hash index hierarchy is used to support CPDP model.

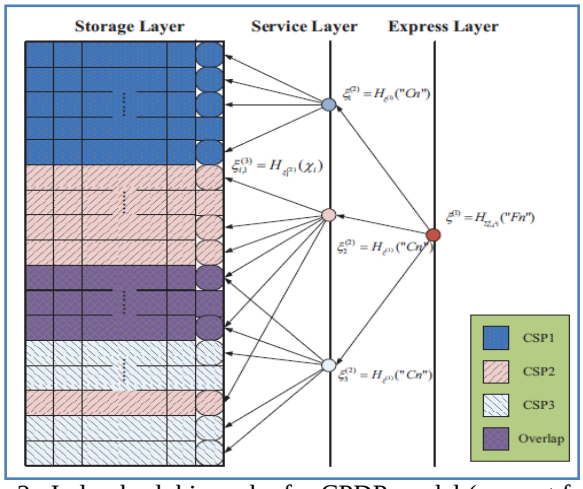

Fig. 2 –Index-hash hierarchy for CPDP model (excerpt from [7])

As can be seen in fig. 2, there are three layers involved in the hierarchy. The express layer is an abstract representation of stored resources. The service layer offers cloud storage services. The storage layer realizes data storage on many physical devices.

IV. CPDP SCHEME

The CPDP protocol is briefly described here. More details can be found in [7]. The scheme has the following mechanisms. They are KeyGen, TagGen, and Proof. The KeyGen is used to generate security keys. It is done by manager. TagGen is meant for splitting file into number of blocks and for each block tag is generated. This tag is used in the data verification phase. The Proof has many steps. They are known as commitment, challenge1, challenge2, response1 and response2. It is described here. First of all, organizer sends commitment to verifier. Then the verifier sends a challenge to organizer. The organizer relays them to the corresponding data blocks. Then organizer gets another challenge from each block. The organizer combines all received responses and the sends it to verifier.

V. EXPERIMENTAL RESULTS

Experiments are made in simulated environment. We built multi-cloud environment by building server programs in Java platform. The programs required by all parties are

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

prepared in Java. The environment used for development is a PC with 4GB RAM, Core 2 dual processor running Windows 7 operating system. We used Net Beans IDE while writing server and client programs. We tested the whole application in a distributed environment. The experimental results are presented as a series of graphs as shown below.

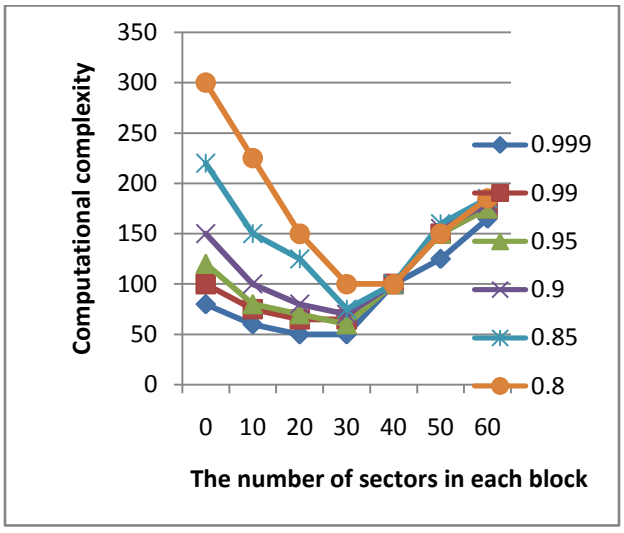

Fig 3 The relationship between computational cost and the number of sectors in each block.

As shown in above figure the horizontal axis represents the number of sectors in each block while vertical axis represents computational complexity

.

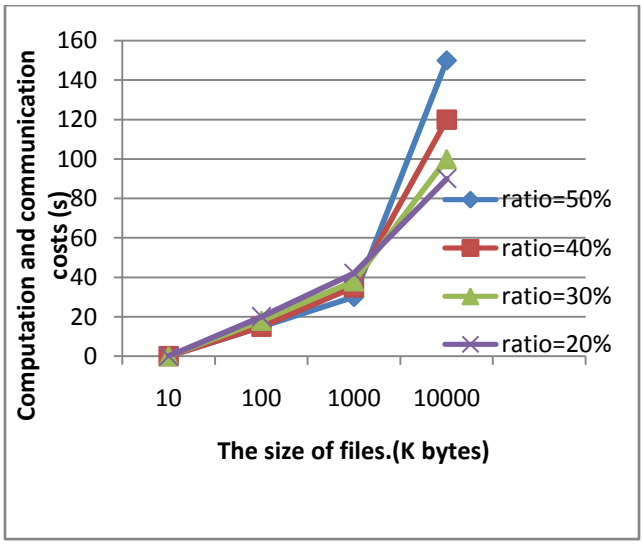

Fig 4 Experimental results under different file size

As shown in above figure the horizontal axis represents the size of files while vertical axis represents computation and communication cost(s).

VI. CONCLUSION

In this paper we implemented the PDP scheme proposed by Zhu et al. [7]. The aim of this scheme is to ensure integrity of outsourced data in multi-cloud environment. To achieve this it makes use of hash index hierarchy and homomorphic verifiable response. The solution is dynamic and scalable in nature. The solution is based on the concept of zero knowledge interactive proof system that can prevent various attacks over cloud. The PDP scheme also has an efficient audit mechanism for periodic verification of data. We built a prototype application that demonstrates the proof of concept. The empirical results reveal that the PDP scheme is very effective and useful.

REFERENCES

[1] A. Juels and B. S. K. Jr., "Pors: proofs of retrievability forlarge files," in *ACMConference on Computer and CommunicationsSecurity*, P. Ning, S. D. C. di Vimercati, and P. F. Syverson, Eds.ACM, 2007, pp. 584–597.

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

[3] C. C. Erway, A. K¨upc¸ ¨u, C. Papamanthou, and R. Tamassia,"Dynamic provable data possession," in *ACM Conference onComputer and Communications Security*, E. Al-Shaer, S. Jha, andA. D. Keromytis, Eds. ACM, 2009, pp. 213–222.

[4] G. Ateniese, R. C. Burns, R. Curtmola, J. Herring, L. Kissner,Z. N. J. Peterson, and D. X. Song, "Provable data possessionat untrusted stores," in *ACM Conference on Computer andCommunications Security*, P. Ning, S. D. C. di Vimercati, andP. F. Syverson, Eds. ACM, 2007, pp. 598–609.

[5] H. Shacham and B. Waters, "Compact proofs of retrievability,"in *ASIACRYPT*, ser. Lecture Notes in Computer Science,J. Pieprzyk, Ed., vol. 5350. Springer, 2008, pp. 90–107.

[6] Q. Wang, C.Wang, J. Li, K. Ren, and W. Lou, "Enabling publicverifiability and data dynamics for storage security in cloudcomputing," in *ESORICS*, ser. Lecture Notes in Computer Science, M. Backes and P. Ning, Eds., vol. 5789. Springer,2009, pp. 355–370.

[7] Yan Zhu, Hongxin Hu, Gail-Joon Ahn and Mengyang Yu, "Cooperative Provable Data Possession for Integrity Verification in Multi-Cloud Storage". IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS.

[8] K. D. Bowers, A. Juels, and A. Oprea, "Hail: a high-availability and integrity layer for cloud storage," in *ACM Conference onComputer and Communications Security*, E. Al-Shaer, S. Jha, and A. D. Keromytis, Eds. ACM, 2009, pp. 187–198.

[9] Y. Dodis, S. P. Vadhan, and D. Wichs, "Proofs of retrievability via hardness amplification," in *TCC*, ser. Lecture Notes in Computer Science, O. Reingold, Ed., vol. 5444. Springer, 2009,pp. 109–127.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 9, September 2013

BIOGRAPIES

Hadassa Katta is student of QIS College of Engineering and Technology, Ongole, AP, INDIA. She has received B.Tech Degree Computer Science and Engineering, and pursuing M.Tech Degree in Computer Science. Her main research interest includes Cloud Computing and data mining.

KOLLA VIVEK received the B. Tech degree from Narayana Engineering College, Gudur in 2009 and M. Tech. degree from VISVODAYA ENGINEERING COLLEGE, KAVALI, in 2012. Currently he is working as an Asst. Professor in QIS COLLEGE OF ENGINEERING & TECHNOLOGY, ONGOLE, PRAKASAM (DT), AP, India. His main research interest includes Cloud Computing and Data Mining

P Raja Rao received the MCA degree from Dr.B.V Raju Institute of Computer Education, Affiliated to Andhra University, Bhimavaram in 2005 .Currently he is working as **Project Lead** in **Polaris Financial Technology** Ltd, Hyderabad, Ranga Reddy dt, AP, India. He is having more than 8 years of experience in Software development and Testing.