International Journal of Computer Applications Technology and Research
Volume 3– Issue 6, 329 - 332, 2014, ISSN: 2319–8656

A Recapitulation of Data Auditing Approaches for Cloud Data

Poonam Dabas
Kurukshetra University
Kurukshetra, India

Divya Wadhwa
Kurukshetra University
Kurukshetra, India

Abstract: Cloud Computing, a buzzword, a technology, has been exploring over the years since 1990s and presently, being considered as a today's s era dependency for the interconnected users. Online resources (like CPU processing power, memory space), software, hardware, capacities are being delivered to the associated users with no headache of handling the intricacies that could come while utilizing the services of this computing methodology. This dependency on server resources for all the major requirements of clients made this computing paradigm, a popular emerging trend in technical world. Now, the biggest concern that one must take into consideration while adopting the benefits of cloud computing is what we call it as 'security'. Obviously, it is the necessary attribute of any computing methodology that is considered as a priority concern. One primary aspect of cloud computing security is maintaining user data integrity which is the key point of this literature review paper. Integrity checking has been designed as a cloud security service by various protocols proposed by many researchers which are being discussed here in this paper.

Keywords: Cloud Computing; IAAS; PAAS; SAAS; Cloud Security; Integrity Checking

1. INTRODUCTION

Cloud Computing has been envisioned as the definite and concerning solution to the rising storage costs of IT Enterprises. Clouds have emerged as a computing infrastructure that enables rapid delivery of computing resources as a utility in a dynamically scalable virtualized manner. Cloud Computing is now on the way. It is a scalable and managed infrastructure and payable as per its usage [1]. The cloud computing technology has been evolved as business cloud models to provide computing infrastructure, data-storage, software applications, programming platforms and hardware as services.

Cloud computing architecture has basically three levels or layers over which it resides and operates. First level is SaaS (Software as a Service). Second layer is PaaS (Platform as a service). The third layer is IaaS (Infrastructure as a service).These levels or layers can also be regarded as cloud service models [2]. A brief idea about these levels is presented below:

- SaaS (Software as a service): At this level, occurrences of a software application can be shared among various users through internet browser. One does not have a need to install and manage particular application, it can be utilized online. Thus, software application is being provided a s a service. Google Docs is a cloud provided SaaS based service.
- PaaS (Platform as a service): In this Layer, customers can develop new applications using APIs which are deployed and configured remotely. There is no need to manage software and corresponding hardware for implementation and placement of application. Google App Engine is one of the PaaS examples.
- IaaS (Infrastructure as a service): In this service model, virtual machines and other abstract hardware and operating system are being made available for cloud users. More clearly, virtualization of computing power, storage and network connectivity is done by this service. The computing resources can be scaled up and down by users dynamically, that is, are hosted by the customer. Example includes Amazon's Elastic Compute Cloud (EC2).

  Thus, the provision of dynamically scalable and often virtualized resources can be made as a service by this evolving cloud computing technology over the internet.

This paper is structured as follows: An introduction to cloud computing and different service models are described in this section. Further, a look is made to cloud security pertaining to data integrity issue through section 2. Next, a review of the related work concerning data integrity checking protocols developed so far is presented in the section 3. Analysis based challenges and future scope is being lightened upon in section 4. Finally, a report of the conclusion of the study of cloud data integrity maintenance is outlined in section 5.

2. A LOOK AT CLOUD SECURITY: CONCERNING DATA INTEGRITY

Security! It is the key for the eminence of a cloud. It is a primary matter of concern for many cloud consumers. Actually, security and privacy interests are a compelling hurdle that is hindering the considerable acceptance of the public cloud across IT entities. As in today's era, computing continuity continues, more devices are adopting more applications, and all these advancements leads to storage of data on external resources which ultimately relates to various privacy issues. These issues comprise the intentional alteration of data without the knowledge of actual data owner, unauthorized access to customer data, not being able to access stored data, unpredicted deletion of crucial information, etc. One of the solution, one might think is to store the data in such a manner that it is of no use to anyone except real owner. Sometimes, the very

much strong concern for cloud security make implementation too complex and bulky, while on the other hand, if more attention is paid over reducing all that complexity, security is compromised as a consequence. Whatever approach is realized, all it needs that all the data protection schemes comprises a compromise between security and the comfort of implementation, no doubt, it should include more secured results [3]. An economic solution is looked upon so that the companies and organizations can focus on their businesses rather than on infrastructure. Well! The very concern of the presented literature review is regarding one aspect of cloud security that is named as Data Integrity Maintenance especially in dynamic cloud environment. Dynamic keyword elaborates in the sense that cloud data does not remain static, it keeps on changing from time to time. Protecting such dynamic data and related alteration comes under this aspect of cloud security. Data integrity is defined as the accuracy and consistency of gathered data, in absence of any modification to the data, while having two continued updates of a file or record. Cloud services should confirm data integrity and provide trust to the user privacy. We had a look over the various methods that have been developed for auditing cloud data and thereby, presenting here in this paper.

3. LITERATURE REVIEW

Various researchers have made major contributions to data integrity maintenance in reference to cloud security access issue. The protocols they developed, utilizes various cryptographic algorithms for implementing and imposing a desired level of security for cloud users so that their data can reside in a secured way, that is, the data could not be seen or altered by some untrusted party on cloud. Zhang Jianhang et al. [4] proposed a new data integrity check scheme based on the well known RSA security assumption. The very obvious advantage of their scheme was that the client did not need to store the copy data in their client side so this indeed freed the client from storage burden. A secure and efficient scheme/ terminology came into play, in which not only data owner but also a third party verifier can check data integrity. Qian Wang et al. [5] proposed an integrated approach of public audibility and dynamic data operations as two salient features of the designed protocol. They concentrated on this very fact because earlier works on ensuring data integrity often lacks the support of either public audibility or dynamic data operations. They designed a protocol that achieved both. A block less approach was adopted and the block tags were authenticated instead of original data blocks in the verification process. This made them noticed. Sravan Kumar R et al. [6] developed an integrity checking scheme which gives a proof of data integrity in the cloud which could be utilized by the customer to check the correctness of data in the cloud. The most important thing that came out of this addressed issue was that storage at the client was kept at minimal that would be beneficial for thin clients. This proposed scheme included an encrypting process that was limited to only fraction of whole data thereby saving on the computational time of the client. But it could not handle the case in which the data changed dynamically. Zhuo Hao, et al. [7] developed a remote data integrity checking mechanism that did not include any third party auditor. Data insertion, modification, and deletion at the block level, and also public verifiability were also promoted by this protocol. The difficulty found was that there was no clear mapping relationship between the data and the tags. Straightaway, data level dynamics could be supported by utilizing block level dynamics. At any time, a piece of data was altered; corresponding blocks and tags were also renewed. Ricardo Neisse et al. [8] presented a system that facilitated periodical and necessity-driven integrity measurements of cloud computing infrastructures. The emphasis was on verifying the integrity of the hardware and software at runtime whenever alterations in the cloud infrastructure were performed. The system developed could remotely keep a check and prove the integrity of necessary system files.

Wenjun Luo et al. [9] addressed a remote data integrity checking protocol based on HLAs and RSA signature with the support public verifiability. Also, this very mechanism was very satisfactory of cloud storage systems because it cloud preserve the file privacy against the third part auditor. This was made possible as the file was encrypted before it was sent to the server and the encryption was kept by the client as a secret. Thereby, the TPA could not be able to get any idea regarding the original file. M. Venkatesh et al. [10**]** proposed an RSA based storage security (RSASS) method which adopted public auditing of the remote data by upgrading existing RSA based signature generation. High level security could be achieved by this public key cryptography technique. The purpose behind using this RSASS method was that, firstly, data storage correctness could be satisfied, secondly, misbehaving server could be determined with a high probability. The preliminary results realized through RSASS, suggested scheme outperforms with upgraded security in data storage when correlated with the existing methods. Henry C. H. et al. [11] studied the problem of remotely checking the integrity of regenerating-coded data against corruptions under a real time operating cloud environment. They enforced the implementation of the functional minimum storage regenerating (FMSR) code and formulated FMSR, which was a code that made the clients to remotely check the integrity of random subsets of long term archival data under a multi-server setting. T.J.SALMA [12**]** explored the problem of data security in a cloud storage, which was actually taken as a distributed storage system. She proposed an effective and competent distributed scheme with a precise dynamic data support, counting block update, delete, and append. This scheme achieved the integration of storage correctness insurance and data error localization, that is, whenever data corruption had been detected during the storage correctness verification across the distributed server(s), mischievous server could be identified simultaneously.

Xiangtao Yan et al. [13] devised a new remote integrity checking scheme for cloud storage that merged correct checking, dynamic update and privacy preserving. In the presented integrity checking scheme, verifier stored only a single cryptographic key and a pre-computed value, irrespective of the size of the file it explores to verify, as well as a small amount of some dynamic state. Yun Yang et al. [14] proposed a fine grained data integrity check scheme. The method compressed effectively the check value for data integrity to reduce storage, and improved the check efficiency of multi-data objects. The focus was made on studying the data integrity of massive storage system in the power cloud computing. Dr. S. Sakthivel et al. [15] enabled a privacy preserving data integrity protection by facilitating public audibility for cloud storage at the hand of third party auditor. The framework used here

was based on an associated PDP protocol that utilizes the challenge – response algorithms and a verification protocol. As third party auditor was included, thus, TPA works on behalf of the data owner who has a huge amount of data to be placed in the cloud. Boyang Wang et al. [16] introduced a novel privacy-preserving public auditing mechanism for shared cloud data, in which, a public verifier could be able to audit the integrity of shared data without having any idea about the private identity information of the group members. The concern was made over the group dynamics. Group dynamics relates to user join and user revocation. Bo Chen Reza Curtmola [17] proposed RDC (Remote Data Checking) scheme that provided robustness and, also, supported dynamic updates, while requiring small, constant, client storage. Remote data checking allowed clients to smoothly test the integrity of data placed at servers that are not trusted. Thus the main challenge that had to overcome was to reduce the client-server communication overhead during updates under an adversarial setting. Yan Zhu et al. [18] introduced a dynamic audit service for verifying the integrity of an un-trusted and outsourced storage. This audit service provided public audibility without downloading raw data and thereby, privacy of the data could be preserved. He Kai et al. [19] proposed a public batch data integrity auditing protocol for multi-cloud storage. Also, fast identification of corrupted data could also be made possible. This could be achieved by making use of homomorphic ciphertext verification and recoverable coding methodology. With this batch auditing methodology adopted in this protocol, the total auditing time could be reduced and the communication cost was also made low.

4. CHALLENGES AND FUTURE SCOPE

As cloud computing is becoming very popular now days. Being a continuing technology in the present era, security aspects of cloud computing is worth concerning. As due to this concern, we have gone through this survey for studying and observing various security challenges. It is a very much fact that the cloud users are often interested in having the integrity of their data stored on the cloud server not get altered in any of the way. The data have to remain intact and not get modified by an unauthorized user. Remote data checking allows for data auditing so that clients could check the integrity of the data at untrusted server. As the task of checking the dynamic data integrity is done by TPA, that is, third party auditor, on behalf of cloud client, the involvement of the client can be eliminated. Moreover, there are a number of challenges in implementing data dynamics. Generating data integrity proofs while considering dynamic nature of the cloud is also contemplated as a challenge for integrity maintenance. Furthermore, block level checking schemes are a bit complex and implementing those in an efficient way can be also regarded as a challenging task. Security and complexity of algorithm are two contradictory terms. A level of balance has to be established between them. Thus, the algorithms being developed for remote data integrity checking should be time and storage efficient and well suited. Thus, Future work aims at implementing these algorithms at minimal costs.

5. CONCLUSION

Cloud computing can be termed as an online service enabling sharing of resources. Proper utilization of the cloud services being offered makes performance of an enterprise well managed in terms of increased efficiency and less overhead incurred. This review paper presented gives us an idea regarding the protocols developed so far, for cloud data integrity maintenance. Many researchers presented their contributions towards this security aspect of cloud computing. They made various successful implementations considering many environmental computing conditions. Thus, very remarkable work has been performed in this area and has made further work more accessible and smoother. We intend to explore this idea of conducting research on this integrity issue.

6. REFERENCES

[1] Wei-Tek Tsai, Xin Sun, Janaka Balasooriya “Service-Oriented Cloud Computing Architecture”, 2010

[2] Sikder Sunbeam Islam, Muhammad Baqer Mollah, Md. Imanul Huq, Md. Aman Ullah “Cloud Computing for Future Generation of Computing Technology”, 2012

[3] Intel IT “Enhancing Cloud Security Using Data Anonymization”, 2012

[4] Zhang Jianhang, chen Hua, “Security Storage in the Cloud Computing: A RSA-based Assumption Data Integrity Check without Original Data”, 143-147 (2010)

[5] Qian Wang, Cong Wang, Kui Ren, Wenjing Lou, Jin Li, “Enabling Public Audibility and Data Dynamics for Storage Security in Cloud Computing”,847-859 (May 2011)

[6] Sravan Kumar R and Ashutosh Saxena, “Data Integrity Proofs in Cloud Storage”, 2011

[7] Zhuo Hao, Sheng Zhong, Member, IEEE, and Nenghai Yu, Member, IEEE**, “**A Privacy-Preserving Remote Data Integrity Checking Protocol with Data Dynamics and Public Verifiability”, September 2011

[8] Ricardo Neisse, Dominik Holling, Alexander Pretschner, “Implementing Trust in Cloud Infrastructures”, 524-533 (2011)

[9] Wenjun Luo, Guojing Bai, “ENSURING THE DATA INTEGRITY IN CLOUD DATA STORAGE”, 240-243 (2011)

[10] M. Venkatesh, M.R. Sumalatha, Mr. C. SelvaKumar, **“**Improving Public Auditability, Data Possession in Data Storage Security for Cloud Computing”, 463-467 (2012)

[11] Henry C. H. Chen and Patrick P. C. Lee**, “**Enabling Data Integrity Protection in Regenerating-Coding-Based Cloud Storage”, 51-60 (2012)

[12] Ms. T.J.SALMA, “A Flexible Distributed Storage Integrity Auditing Mechanism in Cloud Computing”

[13] Xiangtao Yan, Yifa Li, **“**A Wew Remote Data Integrity Checking Scheme for Cloud storage With Privacy Preserving”, 704-708 (2012)

[14] Yun Yang, Lie Wu, Yulin Yan, Cong Xu, “Fine-Grained Data Integrity Check for Power Cloud Computing”, 1346-1350 (2012)

[15] Dr. S. Sakthivel, B. Dhiyanesh, "A PRIVACY-PRESERVING STORAGE SECURITY FOR SPATIAL DATA IN DYNAMICS CLOUD ENVIRONMENT", 2013

[16] Boyang Wang, Hui Li, Ming Li, "Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud", 295-302 (2012)

[17] Bo Chen, Reza Curtmola, **"**Robust Dynamic Provable Data Possession", 2012

[18] Yan Zhu, Gail-Joon Ahn, Hongxin Hu, Stephen S. Yau, Ho G. An, and Chang-Jun Hu, "Dynamic Audit Services for Outsourced Storages in Clouds", April-June 2013

[19] He Kai, Huang Chuanhe, Wang Jinhai, Zhou Hao, Chen Xi, Lu Yilong, Zhang Lianzhen, Wang Bin, "An Efficient Public Batch Auditing Protocol for Data Security in Multi-Cloud Storage", 51-56 (2013)

.