Journal of Asian Multicultural Research for Social Sciences Study
ISSN: 2708-969X
Vol. 3 No. 3, 2022 (page 038-041)
DOI: https://doi.org/10.47616/jamrsss.v3i3.297

Remote Data Auditing in a Cloud Computing Environment

Firas Mohammed Daham[1], Ola Sameer Sabti[1]

Email: Firaspune@gmail.com, ola.samir@duc.edu.iq

[1]Department of Finance and Banking Dijlah University, Baghdad, Iraq

Received: May 28, 2022 Received in Revised: June 21, 2022 Accepted: July 3, 2022

Abstract

In the current paradigms of information technology, cloud computing is the most essential kind of computer service. It satisfies the need for high-volume customers, flexible computing capabilities for a range of applications like as database archiving and business analytics, and the requirement for extra computer resources to provide a financial value for cloud providers. The purpose of this investigation is to assess the viability of doing data audits remotely inside a cloud computing setting. There includes discussion of the theory behind cloud computing and distributed storage systems, as well as the method of remote data auditing. In this research, it is mentioned to safeguard the data that is outsourced and stored in cloud servers. There are four different techniques of remote data auditing procedures that are presented here for distributed cloud services. There are several difficulties associated with data audit methods; however, these difficulties may be overcome by using a variety of techniques, such as the Boneh-Lynn-Shacham signature or the automated blocker protocol. In addition to that, other difficulties associated with distributed-based remote data auditing solutions are discussed. In addition, a variety of approaches might be researched further for further examination in order to find answers to these impending problems.

Keywords: Cloud Computing, Cloud storage providers, Data Auditing

Introduction

The information technology (IT) paradigms of the contemporary period include cloud computing as one of the most essential types of computing services. When there is a requirement to produce economic value for cloud providers out of excess computing resources, it satisfies the desire of users for high capacity, dynamic computing capabilities in diverse utilizations such as data archiving and business intelligence. It also satisfies the need of consumers (Sookhak et al., 2014).

A significant number of users are uploading enormous amounts of data for storage and then deleting their data from servers in their own locations. They download the data they need from a cloud server whenever they need it, but this might open them up to potential security risks, such as assaults by malevolent actors or malfunctions in the underlying hardware or software. The customers of the cloud service provider (CSP) were not informed about the severity of the failures or assaults (Rasheed, 2013). In this scenario, it is necessary to provide evidence in order to demonstrate that the data has been successfully saved on the cloud server. Confirming that data stored on cloud servers is secure is both the most difficult and important challenge presented by cloud computing. Data auditing is a way that allows access to examine the quality of the data in which they may save without accessing it on distant cloud servers. This can be done by accessing the data directly.

Data auditing may be broken down into two categories, private auditing and public auditing, which differentiate themselves from one another dependent on who is doing the verification.

Users are able to verify the accuracy of the data using private auditing techniques; however, this raises the bar for customers, making it unaffordable for them to do so. As a result, public audit techniques enable the audit process to be carried out by any public verifier, and they also provide the consumer with a public key. In most cases, the verification segment is carried out with the assistance of a knowledgeable third-party auditor (TPA) (Sookhak et al., 2014; Rasheed, 2013). In order to determine the accuracy of the data stored in the cloud server, many auditing procedures have been put into place. Nevertheless, in the event that the data is changed, these approaches are not beneficial since they are unable to determine which block is faulty. As a result of the constant requirement for the data to be updated, there is no acceptable authenticated data system that can provide reliable auditing. As a result, the Boneh-Lynn-Shacham (BLS) signature and the automated blocker protocol are both viable solutions for the public auditing system for cloud storage.

Concepts of Cloud Distributed Storage Systems

Distributed Storage Systems

This system stores data across a number of conventional servers, which together function as a single storage system. Additionally, the data is partitioned between these servers. Users are given the ability to remotely store data via the development of distributed storage systems, which also create features such as federation, anonymity, publication, and archiving (Sameen et al., 2021). The most essential argument for utilizing a distributed storage system is that the traditional method of data storage is no longer viable. Despite its high cost and flexibility, the traditional method is not much more efficient than using a distributed storage system.

Remote Data Auditing (RDA)

The auditing of remote data is an important approach that is also highly helpful for auditing the accountability and coherence of external sources of data to spread servers or a single server (Sameen et al., 2021). Remote data auditing is a technique that is significant and extremely beneficial. In order to ensure the long-term dependability of data that has been outsourced to cloud storage providers (CSPs) or to data centers, this kind of assurance is required. The remote data auditing service incorporates a number of protocols to demonstrate the precise distant data that is kept in cloud storage in a manner that is more effective and trustworthy, without the need to download all of the data. In addition, the administration of the outsourced data is carried out by cloud service providers that are not dependable third parties. A spot-checking approach is used in RDA frameworks to validate the data that was outsourced, and this technique needs just a tiny portion of the whole data set, which is required in order for the data set to be considered complete (Prisca et al., 2021).

Methods of RDA Schemes for Distributed Cloud Servies

There are many techniques of RDA schemes to secure the integrity of the outsourced data for distributed storage systems.

Efficient Multi-Copy Provable Data Possession (EMCPDP)

The EMCPDP method is beneficial because it is resistant to attacks by collaborating servers, it does dynamic auditing, and it has less storage overhead than the MRPDP technique. This method is composed of two distinct components, which are known as probabilistic (PEMC-PDP) and deterministic, respectively (DEMC-PDP). The probabilistic approach is similar to the spot-checking method in the sense that it examines a random portion of the file. When using a deterministic technique, each and every file block is checked for accuracy.

Cooperative Provable Data Possession (CPDP)

This technique provides distributed systems with an auditing framework for distant data that is based on replication. It makes use of hash index hierarchy, often known as HIH, in addition to homomorphic verification response (HVR). The structure of HIH is hierarchical, and it displays the relationships that exist within the many data blocks that are maintained by the various storage service providers. It is composed of three layers: the Service Layer, the Storage Layer, and the Express Layer. These layers provide support for batch auditing, which is necessary for auditing dynamic data as well as auditing various clouds.

Tree-Based Dynamic Multi-Copy Provable Data Possession (TB-DMCPDP)

When using the TB-DMCPDP approach, each and every duplicate is given the actual, physical shape of the MHT. When creating an isolated tree, the root of each individual tree is placed in the form of a leaf. Utilizing a directory MHT is the key to achieving the method's overarching goal, which is to validate the authenticity of each and every duplicate in a hierarchical structure. Within this root node are the leaf nodes of the tree, which are copied as MHT for each file.

Map-Based Dynamic Multi-Copy Provable Data Possession (MB-DMCPDP)

In order to discover the storage and improve the computation, a unique data structure known as the map-version table was used, and this was done so as a means of investigation. The serial number (SN), the block number (BN), and the version number are each separated into their own column in the tabular data that is used to assess the externalized data integrity (VN)

Distributed Based RDA Techniques Challenges

Dynamic Data Update

Because data processing in the real world always includes either live data or active log files, RDA's active data updating is one of its most important and distinguishing qualities. In the process of updating using static mode, modifications, deletions, and insertions are all involved. After the preceding tasks have been completed, it is necessary for the users to get the data from the cloud server that was obtained from external sources and then synchronize it. If the audit method includes active updating of data, the user will be asked to download all of the chunks that need to be updated. This will result in a reduction in the amount of calculation performed and an increase in the amount of data exchanged during the process of updating data on the user servers (Arjun & Vinay, 2018).

Batch Auditing

Batch auditing enables TPA to handle a large number of auditing jobs received from different users all at once rather than doing each task separately. This saves TPA time and effort. Because redundancy is one of the properties of RDA algorithms, it is challenging to perform forwarding batch auditing in distributed storage systems (Arjun & Vinay, 2018; Wang, et al., 2012).

Data Deduplication

The data deduplication capability is used to eliminate unnecessary data copies in order to save on storage costs. It is also known as the data compression method that is used to prevent the redundant storage of data (Giuseppe et al., 2013).

Lightweight Data Auditing Approach

The development of a lightweight remote data audit technique helps to increase the security level of mobile users without any kind of constraint, and this demand is an essential problem in the mobile cloud computing environment (Prisca et al., 2021; Giuseppe et al., 2013).

Conclusion

This research touches on the concepts of cloud computing and distributed storage systems, as well as the RDA method. In this research, it is mentioned to safeguard the data that is outsourced and stored in cloud servers. It is explained how to use RDA schemes in four different ways for distributed cloud services. There are many difficulties associated with data audit methods, but these difficulties can be overcome by employing a variety of techniques, such as BLS signature and automatic blocker protocol. In addition to that, certain difficulties associated with distributed-based RDA approaches are outlined. In addition, a variety of approaches might be researched further for further examination in order to find answers to these impending problems.

References

Arjun, U & Vinay, S (2018). A Review on Remote Data Auditing in Cloud Computing. *International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)* 5(4).

Giuseppe A, Alessio B, Walter de D, and Antonio P. (2013). Cloud monitoring: A survey. *Computer Networks 57*(9). 2093–2115.

Prisca I. O., Stanley A.O. and Juliet N. O. (2021). An improved data leakage detection system in a cloud computing environment. *World Journal of Advanced Research and Reviews, 11*(02), 321–328.

Rasheed, H. (2013) Data and infrastructure security auditing in cloud computing environments. *International Journal of Information Management.* JJIM-1286. 5.

Sameen F, Dr. Shafiq H. & Rana A. (2021). An Efficient Secure Auditing Framework for Big Data Storage in Cloud Computing Environment ISSN (2210-142X) *Int. J. Com. Dig. Sys. 10*(1).

Sookhak, M., Gani, A., Talebian , H. , Akhunzada, A., Khan, S., Buyya, R. And Zomaya, A.Y (2014) Remote Data Auditing in Cloud Computing Environments: A Survey, Taxonomy, and Open Issues. *ACM Computing Surveys*.

Wang, C., Wang, Q., Ren, K., Cao, N. and Lou, W. (2012). Toward secure and dependable storage services in cloud computing. *Transactions on Services Computing*, *5*(2) 220–232.