Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 10, 2025

PROVIDING DATA INTEGRITY SERVICE IN CLOUD

Muhammad Hanif[*1], Hazrat Abbas[2], Iqra Bibi[3], Humemon Ghulam Khaliq[4], Anwar Sher[5], Fida Ur Rahman[6]

[*1,2,4,5,6]*Riphah Institute of Informatics, Riphah International University, Pakistan*
[3]*Department of Computer Science, Islamia College and University Peshawar, Pakistan*

[*1]muhammad.hanif@riphah.edu.pk,[2]abbaskhan0345060@gmail.com, [3]iqrakhan191312@gmail.com, [4]aminakhannnn102@gmail.com, [5]anwarsher976@gmail.com, [5]fidaurrahman1416@gmail.com

DOI:https://doi.org/10.5281/zenodo.17263050

Keywords
cloud, data integrity, management services, security, storage.

Article History
Received: 12 July 2025
Accepted: 22 September 2025
Published: 04 October 2025

Corresponding Author: *
Muhammad Hanif

Abstract

The increasing services of cloud is attracting the organizations; thus a lot of new customer is joining the cloud day by day. They not only use the cloud services, but also store their data. Due to cloud characteristics of broad access network. The data can be accessed from anywhere, only require internet. This data now required proper security mechanisms to provide integrity. The data integrity is the main concern of the modern age. Where an unauthorized person cannot change the data. The user data; that is saved on the cloud. Thus cloud service provider can easily access the data and can make changes. In this paper a third party service for data integrity model and its algorithm is presented. So that user can safely put their data on cloud and if any changes occur in that the user get aware of it.

INTRODUCTION

CLOUD computing is the utilizing of resources over the internet. These resources can be in the form of software or hardware. All these resources are provided as a services over the cloud. There are three main service models in cloud computing. They are Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS) and Infrastructure-as-a-Service (IaaS) [1]. These services are presented in the form of stack called cloud stack. The other services fall in these three main categories of the service [2]. In SaaS; the services provider, provides the service of users required software. This service is provide to customers via web interfaces. In PaaS; the service provider, provides the different types of development and database platforms to the customers. While in IaaS; the service providers, provides the hardware and computing services to the customers. These services are in the form of virtual machines. The emerging services of cloud is attracting new customer day by day. This new services is now refer as "Every-thing As a Service" or "Any-thing As a Services. This is denoted by XaaS. In this terminology each and every thing a service provider can provider occur. According to NIST (Nation Institute of Standard & Technology, there are four deployment models in cloud [2]. They are Public Cloud, Private Cloud, Community Cloud and Hybrid Cloud. Depend on the customer and users which cloud is suitable for them. In Public Cloud all the services

provided by a service provider can be publically access. This is open cloud everyone can use and the services may be free of cost or paid. The Private Cloud is basically the cloud services provided by an organization to its customers or employees. These services are restricted only to that organization. No one can access them from outside the organization. And only authorized persons of that organization can access and use those services. On the other side when multiple same type organization combine and develop a particular cloud that can only be used by those organization employees, such cloud is refer as Community Cloud. The services offer via this type of cloud is only bounded to those particular organizations for which that cloud is developed. The Hybrid Cloud is the type of cloud, where more than one type of clouds combine to share their services. Such as when a private cloud is access via public cloud. The hybrid cloud provide facility to private cloud or community cloud to extend their boundaries. Along with these models there are five main characteristics of cloud. They are "On demand self-service" which mean that each and everything in the cloud is automatic. When a customer request for a service that service will be provided to customer automatically without any administrator or manual control. "Broad access network" mean that cloud services can only require network access. And these services can be access from anywhere. "Resource pooling" means that properly utilize the resources are available to the customers. And if a customer is not using a resource that will be assign to another customer. "Rapid elasticity" means that scalability is available in the cloud. The services can be extend for the customers when their demand extend. And at last "Measured service" means that for each service in the cloud there will be unit in which those services will be measured. The cloud is based on the "Pay as you go" option. Thus how much the customers will use the services that much payment will come. This unit depend on the services provider what they suggest. It might be time period, or it might be data size, or functions used in the services. The unit can be anything that is mention in the service level agreement (SLA) between customer and the service provider.

As the number of customer is increasing each day. The more data is uploading over the cloud. This data can be in any format. It is in the form of images, text, audio and / or video. In cloud computing the main concern is the security. The security attacks can be either external or internal. The external attack mean that attacker get or access the data outside the cloud. While in internal attack the attacker accesses the data inside the cloud. These attacks can of either passive or active. In passive attack, the attacker only accesses the data and made no changes in the data. While in active attack, the attacker not only access the data but also made changes in the data. The main goals of security are confidentiality, integrity and availability. In passive attack the confidentiality of data is suffer. In which the attacker (unauthorized person) access the data. While due to active attack the integrity of data is suffer. In which the attacker (unauthorized person) not only access the data but also made changes in the data. The third goal of security is availability, which is effected due to the Denial of Service (DOS) attacker. In DOS the attacker aim is to stop access of authorized person to data or service [3].

Integrity of data is the main concern of today's cloud. In integrity only authorized person can change the data. Or an unauthorized person cannot make changes and modification in the data. The data integrity is required at all the three stages in the cloud. During the uploading of data to the cloud the data must be consistent. Then during in storage stage the data must be safe from modification and changes. And finally during downloading the data must be correct. The integrity can be achieved through hashing algorithm [4]. Currently hashing algorithms like MD5, SHA1, SHA2, etc. provide basic hashing facility to the customers. Such algorithm works on simple principles of finding the hash value of the data. After passing the data through a particular hashing algorithm, a hash value for that data is generated. The hash value is also called the finger print of the data. This hash value will remain same for only this particular data. Even for a very minute changes or single bit change in the data will provide different hash value. The working mechanism of hashing algorithm is quite simple. It get a particular size of data as an input an produce a fix length hash value. The nature of hash value is iterative. It split the huge data into smaller fix length parts and pass through the hash function. The hash function is the heart of the hashing algorithm [5]. It takes two input. One the data part and other the previous data part hash value. For

the first data part the hash value is by default provided by the algorithm. And thus for each and every part the hash value is generated which work is the second input for the next data part. The final value generated by the algorithm is the hash value of the entire data. This final value is called the message digest of given input data. In hashing the collision can also be never under estimate. The hashing collision mean that for two different data set a same value of hash is generated. For a best hashing algorithm the collision is not the affordable. But at the same time it cannot be ignored.

In cloud environment the integrity is not yet implemented [6]. Thus data on the cloud can be easily modified and changed. This modification can be done by external attacker or internal attacker. The change in data can be made in either stage. It can be made during uploading, storing or downloading stage. Thus this paper provide the third party services for data integrity which can be implemented to get the best results. A data integrity scheme is presented in this paper. In which the hash value of data is generated on the source side and both on the third party integrity provider store side called Integrity-Management-System (IMS). The hash value is kept on the IMS. Whenever the data is downloaded by any user the hash value will be generated by that user and will be matched with the hash value on the IMS. If both are matches the data is safe and correct. If not there is change in the data. The rest of the paper is distributed as following. In section II literature review is presented. In section III proposed methodology is presented and results and finally the conclusion is presented in the last of paper.

Literature Review

In cloud, the data integrity is always the main concern. Large amount of work is so for done in variety of areas. These areas include from storage to architecture and then to standard and still in progress. In cloud the main issue with integrity is that it is always treating along with other security aspects. The data integrity require a separate portion of work. To isolated deal with its problems and issues to provide best result [7].

In today's world data is the heart of the technology. The main war and struggle around the globe is over data. Thus data need to be correct and constant. It is extremely important that at any level, the cloud services provider must assure the data integrity to their customer whenever it is demanded. This is the assumption in people that data is always safe and sound when it is uploaded and stored. But in current era this assumption is not true [8]. Variety of attacks has been detected over the integrity during its storage and transmission. Over the data integrity storage mechanisms the work is done in [9]. Not only in cloud storage, is the mentioned work implemented but also in data centers and over local servers. The integrity has three main techniques, Checksum, Mirroring, and RAID [10]. The Checksum is to finding the hash value of the data via different algorithms like MD4, MD5, SHA1, SH2. In this paper our algorithm and scheme works on this mention technique. In mirroring the two or more sets of copies are stored. The integrity is achieved by comparing the data sets with each other. In RAID the physical component of storage are virtually integrated to check the integrity of the data through comparison.

In [11], Using RAID technique for data integrity was provided. But this technique is quite slow due to binary data conversion and computation. To measure the performance of different encryption algorithm the work was presented in [12]. These algorithms was tested over the text files and using DES, AES and RSA. The computation time, output results and buffer used was the parameters. The result was shown that in computation RSA took more time than other. The data integrity provided via matrices was proposed in the previously mentioned paper. To ensure the integrity the determinant of each data matrix is generated. The model works on double algorithms, i.e. AES and SHA1. And along this digital signatures were used for data security. This approach was quite interesting but time consuming. In [13], they presented a third party authentication scheme. This scheme was to achieve the privacy of data in public cloud. This scheme also provided help in audits in cloud storage environments.

For data integrity in cloud environment work is still in progress [14]. Providing data integrity services via comparison with original data without downloading is done. But this solution is suitable in static environment. This operation not support the dynamic data operation. Along this [15] proposed a solution to determine the data integrity by comparing the data file with server file. They also researched on

the dynamic data operation. They provide provable modification in the stored data. But the provided algorithm working was very slow. In cloud, few standards were used to deal with integrity issues. In DBM, the data integrity is obtained through ACID. The standard name SNIA was developed to provide data protection in cloud storage.

For third party integrity, the DIaaS was proposed by [16]. They proposed third party data integrity services to their customers. The IMS was used to low down the burden form the cloud service provider. Where the information about the data store and thus on demand those information is provided to the customers to check their data either corrupted or not. [17] also presented the idea of third party data integrity. In dynamic cloud storage they verify the cloud customer request that either changes had done or not in their data. There are variety of work done on data privacy. The use key-management-system for solving issues in data integrity is one of the simplest method follow by many researchers. For data privacy in public cloud [18] proposed a virtual file store to achieve the integrity.

Providing data integrity via ID based analysis was proposed by [19]. For remote storage and cloud environment they combine the PKI and RAID schemes to achieve the data integrity. A public integrity system was proposed by [20]. Which was based on multi copies and hash tree. Their aim was to propose a method for privacy, to be maintained in public cloud. The privacy and remote data checking modification protocol was designed and their aim was to predict the change and modification when ever done in the already stored data on public cloud. The work done on data outflow by [21]. They identified and secure data leakage to maintain the integrity using a dynamic key protocol. The purpose of this dynamic key protocol was for secure and correct data transmission not to violate the contents of the data. To provide multicopy protection of data over the public cloud, and also for the auditing in the cloud. The auditing in cloud is the main concern in the public cloud. They presented it by hiding the keys through which the encryption is done. Thus Providing data privacy over public cloud.

Methodology

The proposed system comprises of four components that works on an algorithm. Fig. 1 shows the working process of the proposed system. The system is composed of mainly four components. The user data, hashing algorithm, Integrity Management System (IMS), and cloud. The user can be both source and destination. They are also called sender and receiver. The sender are those who has the data while the receiver are those who required the data. The working condition is mentioned in the following proposed algorithm.

Proposed Scheme for data integrity

1. The source will first take the hash value of the data. The hashing algorithms is defined and the hash of the data is generated.
2. Source upload the data to IMS. Plus request for the hash value of the data over that algorithm which the source had already used.
3. The IMS received the data and stored it temporary. The hash value is generated and sent it to the source.
4. Source received the hash value from IMS and compare it with its previous generated value. If both matches the hash value is then digitally signed by the source and transmit this certificate to IMS.
5. IMS received the digitally signed value and verify the signature and also signed it and store it. IMS save the data on the cloud.
6. Whenever destination download the data from cloud. The destination will first generate the hash value of the data.
7. Destination asked for the digitally signed hash value certificate from the IMS.
8. The IMS will transmit the hash value of the data.

Destination compare both the hash values. If both matched the data is correct and no modification had done.

TABLE I
NOTATIONS FOR AGORITHM

Symbol	Quantity
SU	: Source User
D	:Data
DS	:Digitally Signed
H	:Hash Algorithm
HV	:Hash Value
IMS	:IntegrityManagementSystem
C	:Cloud
DU	:Destination User
CM	:Compare
T	:Temporary

Proposed Algorithm for data integrity

SU :H(D) = HV
SU→IMS :T(D), H(D) = HV
SU :CM(HV)
SU→IMS :DS(HV)
IMS :C(D)
DU :D
DU→IMS :HV
IMS→DU :HV
DU :CM(HV)

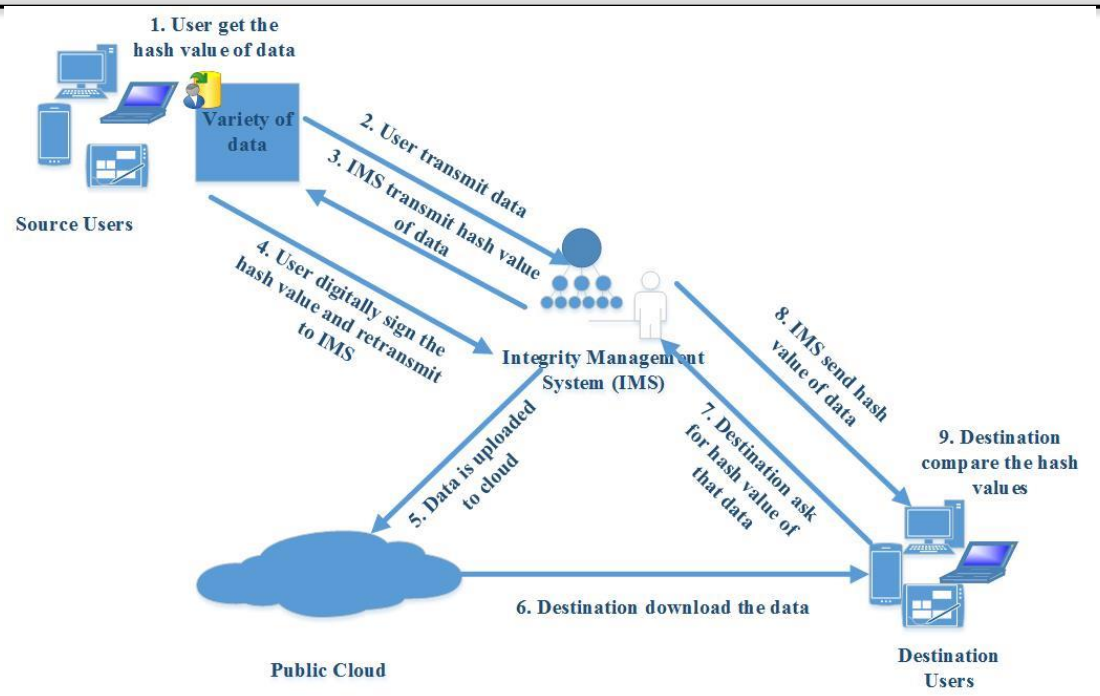

Fig. 1.Proposed Data integrity Scheme.

IMS→DU :HV
DU :CM(HV)

IV. Working Of Algorithm

The working structure can be discuss as of the proposed methodology as mentioned in the algorithm. The User (source) will first of all having the data will generate the hash value with a particular hashing algorithm and transmit the data via data integrity application to IMS. The IMS received the data and the mentioned hashing algorithm. Via that mentioned algorithm the IMS also generate the hash value and transmit it to back to the source side. The time taken in generating a hash value depends on the processing power and hashing algorithm. For a core i7 with 2.7 GHz octacore, and MD5 hashing algorithm the following table II and graph is generated.

The source received the hash value and compare it with his already generated hash value. If both matches it means that the data is safely transmitted to the IMS. If not matches it mean that some changes and modification had occurred in the data. Therefor in case of different value the data will be canceled and all the above operations will be repeated again until same hash value obtained. In case of same hash value the source digitally signed the hash value and upload that as a digital certificate for that particular data to IMS as shown in Table 3 and Fig. 3.

TABLE III
DIGITAL SIGNATURE TIME

File Size (KBs)	Time (Seconds)
10	5
20	9
30	16
40	24
50	28
60	35
70	40
80	45
50	6
60	7
70	8
80	9
90	10
100	11

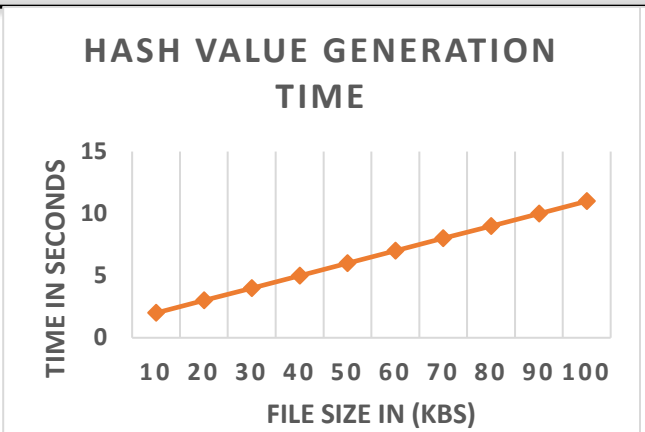

Fig. 2. Encryption Time Graph.

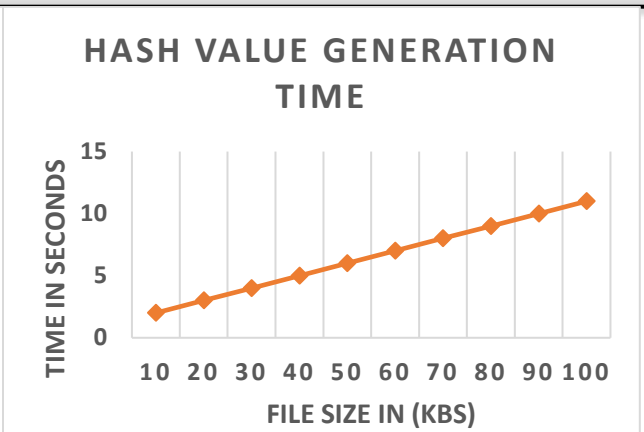

Fig. 3. Digital Signature Time Graph.

The IMS upload the data to the cloud and save its digital certificate with itself. Whenever a user (destination) download the data. That destination will also request for the digital certificate of that particular data from IMS. The IMS will not only provide its certificate but will also mention the hash algorithm via which this value is generated. The destination will then generate the hash value of that downloaded data and compare it with the digital certificate hash value. If both matches it mean that the data downloaded by the destination is the same as was uploaded by the source side with no changes and modification.

Conclusion

This paper presents a new technique for data integrity. The idea behind this was to provide a third party service for data integrity. So that the current services and operational time not suffer. In proposed technique, the hash value of data is generated by the user and then data is transmitted to the IMS. The IMS also generate the hash value and forward it to the user the user compare both the hash values and when matches signed it and upload it to the IMS. The data is also uploaded to the cloud and thus when the destination download the data. The destination request for the source signed hash value and matched it with his generated value when both matches the data is secured and no changes had made. This integrity service will enhanced the protection of data from unauthorized changes and modification and will improve the cloud services.

References

[1] V. K. Gali, S. S. Chauhan, G. K. Gupta, P. R. Bulani, V. Y. Ravalji, and K. D. Jayaraman, "Ensuring Data Integrity in Multi-Tenant Cloud Environments Through Advanced Testing," in *2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT)*, 2025, pp. 1365–1369.

[2] P. Mell and T. Grance, "The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology," *Nist Spec. Publ.*, vol. 145, p. 7, 2011, doi: 10.1136/emj.2010.096966.

[3] W. Zhang and L. Chen, "Developing a Zero-Trust Security Model for Cloud Migration: Ensuring Data Integrity and Confidentiality in Hybrid Cloud Architectures," *Adv. Theor. Comput. Algorithmic Found. Emerg. Paradig.*, vol. 15, no. 2, pp. 15–27, 2025.

[4] E. Alaka, K. Abiodun, S. O. Jinadu, E. Igba, and V. N. Ezeh, "Data Integrity in Decentralized Financial Systems: A Model for Auditable, Automated Reconciliation Using Blockchain and AI," *Int. J. Manag. Commer. Innov.*, vol. 13, no. 1, pp. 136–158, 2025.

[5] E. Uzoma, J. O. Enyejo, and T. Motilola Olola, "A Comprehensive Review of Multi-Cloud Distributed Ledger Integration for Enhancing Data Integrity and Transactional Security," *Int. J. Innov. Sci. Res. Technol.*, vol. 10, no. 3, pp.

1953–1970, 2025.
[6] N. L. Khumalo, T. E. Mathonsi, and S. O. Ojo, "Implementation of an Enhanced Data Integrity Verification Model in Cloud Computing," in *2024 International Conference on Electrical, Computer and Energy Technologies (ICECET*, 2024, pp. 1–6.
[7] Y. Liu, "The Latest Application and Security Analysis of Cryptography in Cloud Storage Data Audit [J]," *Procedia Comput. Sci.*, vol. 259, pp. 984–990, 2025.
[8] Y. Zhao, Y. Qu, Y. Xiang, M. P. Uddin, D. Peng, and L. Gao, "A comprehensive survey on edge data integrity verification: Fundamentals and future trends," *ACM Comput. Surv.*, vol. 57, no. 1, pp. 1–34, 2024.
[9] X. Wang, "Research on cloud storage data integrity verification protocol based on smart technologies," in *International Conference on Cloud Computing, Performance Computing, and Deep Learning (CCPCDL 2024)*, 2024, p. 1328102.
[10] W. Li, W. Susilo, C. Xia, L. Huang, F. Guo, and T. Wang, "Secure data integrity check based on verified public key encryption with equality test for multi-cloud storage," *IEEE Trans. dependable Secur. Comput.*, vol. 21, no. 6, pp. 5359–5373, 2024.
[11] M. K. Chandol and M. Kameswara Rao, "Blockchain-based cryptographic approach for privacy enabled data integrity model for IoT healthcare," *J. Exp. \& Theor. Artif. Intell.*, vol. 37, no. 1, pp. 53–74, 2025.
[12] P. Goswami, N. Faujdar, S. Debnath, A. K. Khan, and G. Singh, "Investigation on storage level data integrity strategies in cloud computing: classification, security obstructions, challenges and vulnerability," *J. Cloud Comput.*, vol. 13, no. 1, p. 45, 2024.
[13] W. Cao, W. Shen, J. Qin, and H. Lin, "Privacy-preserving and verifiable convolution neural network inference and training in cloud computing," *Futur. Gener. Comput. Syst.*, vol. 164, p. 107560, 2025.
[14] J. Lapmoon and S. Fugkeaw, "A Verifiable and Secure Industrial IoT Data Deduplication Scheme With Real-Time Data Integrity Checking in Fog-Assisted Cloud Environments," *IEEE Access*, 2025.
[15] C. Yang, Y. Liu, Y. Ding, and H. Liang, "Secure data migration from fair contract signing and efficient data integrity auditing in cloud storage," *J. Netw. Comput. Appl.*, vol. 239, p. 104173, 2025.
[16] M. Z. Hasan, M. Z. Hussain, Z. Mubarak, A. A. Siddiqui, A. M. Qureshi, and I. Ismail, "Data security and integrity in cloud computing," in *2023 international conference for advancement in technology (ICONAT)*, 2023, pp. 1–5.
[17] L. L. Scientific, "Data integrity concerns requirements and proofing in cloud computing," *J. Theor. Appl. Inf. Technol.*, vol. 102, no. 12, 2024.
[18] B. Hazela, S. K. Gupta, N. Soni, and C. N. Saranya, "Securing the confidentiality and integrity of cloud computing data," *ECS Trans.*, vol. 107, no. 1, p. 2651, 2022.
[19] R. Sharma, V. Sharma, T. K. Vashishth, B. Kumar, B. K. Sharma, and S. Chaudhary, "Implementation of adaptive cloud security framework for cloud computing monitoring system applications," in *AIP Conference Proceedings*, 2025, p. 20014.
[20] N. L. Khumalo, T. E. Mathonsi, and S. O. Ojo, "An Improved Data Integrity Verification Model in Cloud Computing," *Eng. Innov.*, vol. 14, pp. 117–125, 2025.
[21] M. S. Mushtaq, M. Y. Mushtaq, M. W. Iqbal, and S. A. Hussain, "Security, integrity, and privacy of cloud computing and big data," in *Security and privacy trends in cloud computing and big data*, CRC Press, 2022, pp. 19–51.