iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** I **Month of publication:** January 2019

DOI: http://doi.org/10.22214/ijraset.2019.1129

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue I, Jan 2019- Available at www.ijraset.com

Meta Data Based Data Integrity Tracking Scheme in Cloud Environment

Patil N.N. [1], Mapari R.B.[2]
[1, 2]*Department of Computer Science and Engineering, Maharashtra Institute of Technology, Aurangabad*

Abstract: Cloud computing is a recent trend in IT in which data and computation moves from local devices to remotely located large scale processing hubs and storages.
It can be an effective solution for enterprises which saves the rising cost of IT setups and management. However, in Cloud platforms user's data is moved into remotely located servers such that users lose control over the data. Along with many advantages, there are many security and privacy related challenges that cloud is facing.
These challenges are need to be clearly understood and resolved. Proof of data integrity means correctness of user data in cloud storage is one of the important concerns that need to be addressed.
Users can not access data stored in the cloud physically. Therefore, a technique or mechanism is required where user can check the correctness of data which is stored in the cloud.
This paper address the problem of data integrity by proposing a mechanism. Through this mechanism users can check the correctness of their data which they have stored on cloud servers.
For this purpose we studied different schemes which had already proposed and found that there is scope of improvement in existing schemes.
So by doing some modifications we proposed a scheme which is less complex and can give better results than existing schemes.
Keywords: Cloud computing, data integrity, Cloud storage, TPA, data outsourcing, untrusted storage.

I. INTRODUCTION

The rapid growth of information and communication technology over the past decades led to the vision that computing will be the essential part of life.
Like all other existing services in daily life computing consist of commodity services. End user can avail those services at any time (on-demand) without necessity of any hardware or software infrastructure. User access the services as per requirement without any concern of how they are getting delivered and from where they are getting delivered.
In cloud data is geographically dispersed and which is accessed by number of users simultaneously and independently through internet. In such a shared and distributed environment, data transfer is done though communication networks. Number of transactions increases as the number of user and amount of data increases. With the dispersed data, significant data transfer increases the chances of data lost, unauthorised access and data alteration. This is one of the fundamental concerns to ensure the data integrity and security in cloud environment. In this environment data owners are not responsible for managing the data but loose the physical possession of their data.
In such a environment it's very challenging to maintain data integrity.
As per our research the problem of data integrity proofs in cloud environment has not been investigated widely and still it's in early stages. Number of existing approaches are either rely on third party services or storing the same data on both the side to check the integrity. We need a comprehensive solution to facilitate cloud users easy way for checking data integrity by combining several different techniques which rely on local resources of the cloud environment. For this purpose we are using a powerful local resource called Metadata means data about the data.

II. BASICS OF CLOUD COMPUTING

Cloud computing can be defined as a computing service which is delivered through the internet rather than physically having computing resources at the end user location. With increase in end users data storage is prioritised in all fields. Now a days businesses are spending huge amount to store and maintain their data. All businesses cannot afford in-house IT setup and support. In such a cases cloud computing is a cheaper solution. Perhaps its advantages and features have succeeded to attract even bigger business as well.

A. Benefits of Cloud computing

Many organisations have adopted cloud and the reason is cloud computing gives freedom to end users by offering services as per their requirements and pay as per use.

Due to cloud computing it is possible to run IT operations without much resources.

Following are the benefits of cloud computing: Improved performance, Low maintenance backup and recovery, Scalability, increased storage capacity, instant software updates etc.

B. Types of Clouds

There are four different types of cloud models that you can use according to requirement

1) *Private cloud:* In this computing resources are deployed for a particular organization. This model is used in intra business interactions.
2) *Public Cloud:* This type of model is used usually for Business to consumer type interactions.
3) *Community Cloud:* This type of model is used for community and organizations.
4) *Hybrid Cloud:* This type of cloud can be used for both type of interactions that is business to business and business to consumer.

C. Cloud Computing Services

There are three different types of services that cloud computing offers

1) *SaaS (Software as a Service):* It is a software distribution model in which applications are hosted by a cloud service provider and made available to end user over a network (internet). Generally we have to purchase software then installed it onto our computer. On the other hand SaaS users can subscribe it usually on monthly basis and use it via internet.
2) *PaaS (Platform as a Service)*: It provides a platform and environment to allow developers to develop or build applications and different services. The service which is hosted by cloud can be accessed through internet.
3) *IaaS (Infrastructure as a Service):* It includes computing infrastructure like networking, server space, bandwidth etc which is virtualized.

III. SECURITY AND PRIVACY ISSUE IN CLOUD

In this section we will see the various security and privacy issues that cloud is dealing with as follows

A. Data breach

As cloud contains massive data into it, cloud service providers are the attractive target for attackers. Attack severity is depend upon the confidentiality of data which will be hacked. Data related to financials can cause severe damage.

B. Network security

As in cloud computing services are delivered through internet, strong network security will be required to avoid leakage of the confidential data.

C. Malicious insiders

Malicious insider means a person who is a part of organisation and has accreditations to get organisation information.

Malicious insiders also can harm more than any other can do.

D. Data integrity proofs

Data integrity means correctness of user data. A mechanism is needed where user can check the integrity of the data stored in the cloud. So in next section we will discuss about the various techniques or schemes introduced by different authors and their limitations.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue I, Jan 2019- Available at www.ijraset.com

IV. LITERATURE REVIEW

As shown in table below, there are number of schemes proposed to overcome the issue of data integrity check in cloud environment. We also mentioned their advantages and limitations.

TABLE 1
Literature Review

Journal Title Schemes	Authors	Concept	Advantages	Limitation
Provable data possession [2]	G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song,R.	Hash function is used	Strong data integrity verification, Reduced network trafic	High computational cost on the server end for computing hash value of each file
PORs: Proofs of retrievability for large files [3]	Juels A, Kaliski Jr BS.	Cryptographic techniques such as sentinel based Values are used	Less storage overhead on server side because storing only sentinel values	Computational overhead for pre-processing the sentinel value
Third party auditor	There are number of Authors who proposed different schemes by using TPA	Key generations based scheme utilized by third party	TPA can apply multiple integrity schemes for checking data integrity, Users can choose different TPA based on their preferences	Retrieval of data blocks from third party for checking data integrity causes privacy issues, More cost due to involvement of third party auditors.

As shown in Table 1, there are number of schemes have been proposed and each scheme has its own advantages and limitations. We found that some major drawbacks like high response time, storage overheads and computational overheads. Also we can overcome those limitations by using different approaches. In upcoming sections we will discuss our proposed scheme in which we tried to overcome one of the limitations that is response time.

V. PROPOSED SCHEME

As discussed earlier the objective of our proposed scheme is to reduce the response time for integrity check by eliminating third party auditor and storing the metadata in the database. In literature review we have seen that there are number of schemes have been proposed to check data integrity in cloud environment by using third party auditor and key generation algorithm like scheme proposed by Yong Yu, Yannan Li, Bo Yang and Willy Susilo which is Attribute based cloud data integrity auditing scheme[8]. Unlike this scheme we are elimination third party auditor, key generation algorithm and using only metadata to prove the integrity of the data stored on cloud.

A. Proposed work

The scheme that we have proposed is very simple, easy to understand and easy to implement. We are simply eliminating the factors which make the process complex and increase the response time. The factors which we are eliminating are third party auditor, key generation algorithm etc. Additionally we are using metadata to prove data integrity and which will be stored in database rather than xml file. Each and every action which is done on data item will be tracked and by using that information we will identify whether the integrity is violated or intact. Fig. 1 shows the block diagram for our proposed scheme which consist of 4 stages as show in figure.

User upload the file while uploading the file, metadata related to data item captured and get stored in the database (table 1). Once file get uploaded it will again capture the metadata and stored in database (table 2). Further any modification on data file will be captured in the database. For integrity checking on user request both the tables will be compared and if any mismatch found proves that data has been tempered.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue I, Jan 2019- Available at www.ijraset.com

B. Architecture of proposed scheme

Fig. 2 shows the architecture of proposed scheme. As shown in figure we are using two cloud instances here, first one is for all computations and second for data storage. Cloud instant 1 will be responsible for storing metadata into the database and for all computations. Cloud instant 2 will be responsible for storing data or works as a cloud storage. End user is accessing the cloud services through access point.

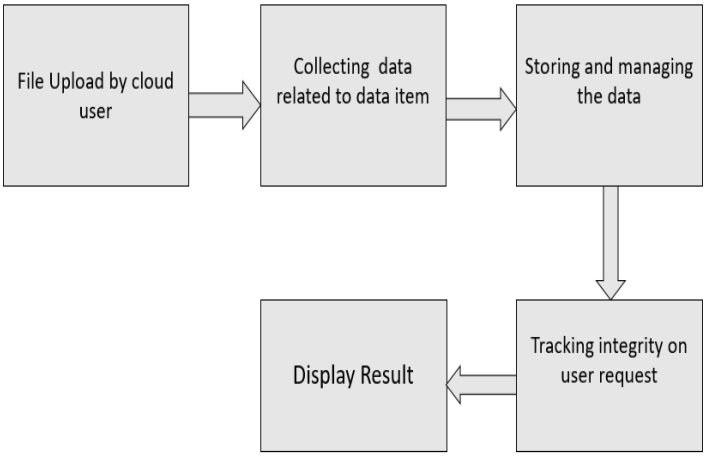


Fig. 1 Bock diagram proposed scheme

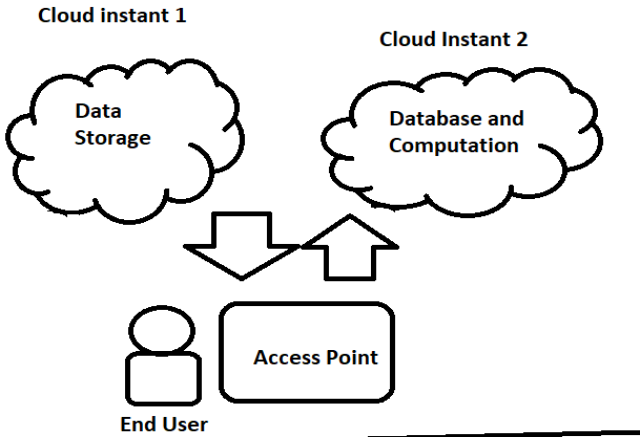


Fig. 2 Architecture of proposed scheme

C. Implementation

In this section we are discussing about the implementation of our idea in the form of working system. We have considered different modules as per requirement of our scheme. Each module is working separately by taking input from previous stage and generate output by using data generated by previous stage. We have used java programming language for development and Mysql database for storing the results. Also user two intendances of clouds one is for data storage and another is for computations.

D. Performance Analysis

This section shows the results that came out after running the system. The performance of the proposed scheme is evaluated. In our implementation, all the algorithms are conducted on a Win 10 64-bit laptop with Intel Core (TM) i5-8250U @ 1.60GHz CPU. The projects are written in java language and java server pages are used. The system is thoroughly tested on both online and offline environment to evaluate the performance. File set ranging from 10 KB to 50 KB is taken as a data set. Response time measured in seconds. The performance of proposed technique is shown in terms of response time. The response time is calculated as below

Response Time = Time taken by system to display the result once query is submitted

Table 2
Proposed scheme Results

Size in KB	Online (in Sec)	Offline (in Sec)
10	0.3	0.1
20	0.53	0.18
30	0.87	0.29
40	1.3	0.37
50	1.7	0.48

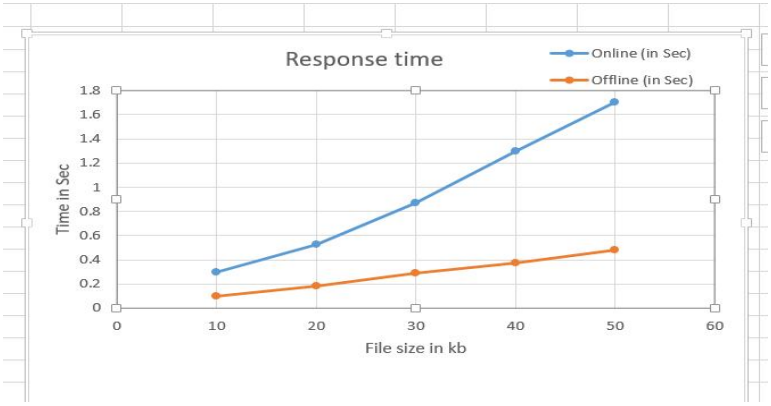


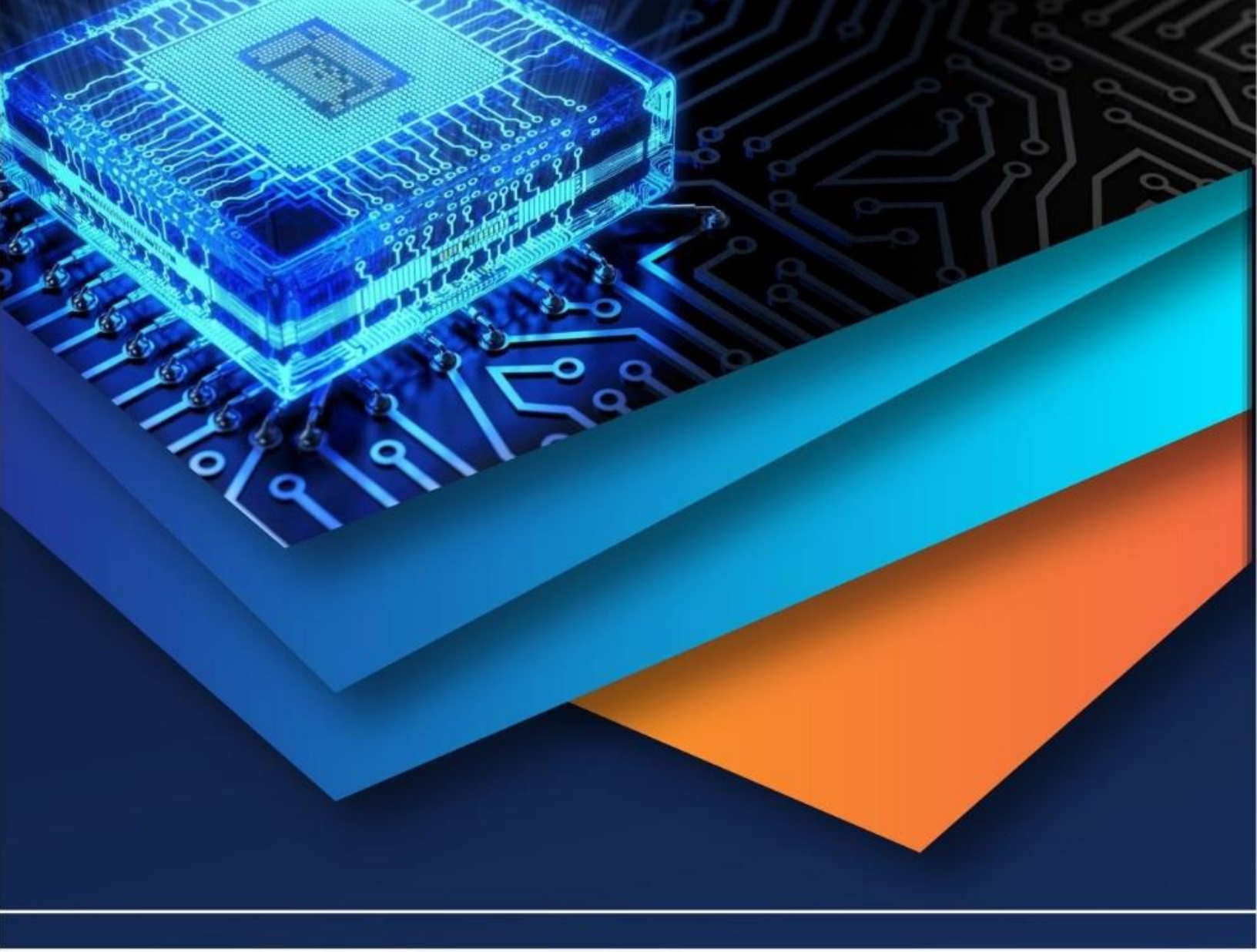
Fig 3 Results Graph

VI. CONCLUSION

In this paper we discussed various security and privacy issues in cloud computing. We found data integrity is one of the important concern that need to be clearly understood and resolved. We studied different integrity checking schemes, their advantages and limitations. We also analysed some existing schemes and identified some issues in them. We found response time is one of the important issues that need to be rectified. So we proposed metadata based data integrity proving scheme in cloud environment. In this scheme we used only metadata to check the integrity of the file. Through this scheme we successfully reduce the response time of integrity checking. The successful execution of our scheme without using any additional entity like TPA clearly proves the Utility of our scheme in cloud environment.

REFERENCES

[1] Daliya Attas , Omar Batrafi “Efficient integrity checking technique for securing client data in cloud computing”, IJECS –IJENS Vol:11 No:05,2011
[2] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, “Provable data possession at untrusted stores,” in CCS ’07: Proceedings of the 14th ACM conference on Computer and communications security. New York, NY, USA: ACM, 2007, pp. 598–609.
[3] A. Juels and B.S. Kaliski Jr., “Pors: Proofs of Retrievability for Large Files,” Proc. 14th ACM Conf. Computer and Comm. Security (CCS ’07), pp. 584-597, 2007.
[4] Amazon.com. Amazon simple storage service (Amazon S3), 2007.
[5] Sravan Kumar R, Ashutosh Saxena,”Data Integrity Proofs in Cloud Storage,”ISBN: 978-1-4244-8953-4/11/$26.00 c 2011 IEEE
[6] Satyakshma Rawat, Richa Chowdhary, Dr. Abhay Bansal,” Data integrity of cloud data storage (CDSs) in cloud” ijarcsse Vol. 3, Issue 3, March 2013.
[7] Saranya Eswaran, Dr. Sunitha Abburu “Identifying Data integrity in cloud storage”, IJCSI Vol. 9, Issue 2, No 1, March 2012.
[8] Yong Yu, Yannan Li, Bo Yang, Willy Susilo,” Attribute-Based Cloud Data Integrity Auditing for Secure Outsourced Storage”, IEEE, VOL. 14, NO. 8, AUGUST 2015
[9] E. Mykletun, M. Narasimha, and G. Tsudik, “Authentication and integrity in outsourced databases,” Trans. Storage, vol. 2, no. 2, pp. 107–138, 2006.
[10] T S Khatri, Prof G B Jethava,” Survey on Integrity Approaches used in the Cloud Computing”, IJERT Vol. 1 Issue 9, November-2012.
[11] C. Erway, A. Kupcu, C. Papamanthou, and R. Tamassia,“Dynamic Provable Data Possession,” Proc. 16th ACM Conf. Computer andComm. Security (CCS ’09), 2009.
[12] Feifei Liu, Dawu Gu, Haining Lu,” An Improved Dynamic Provable Data Possession Model,” 978-1-61284-204-2/11/$26.00©2011 IEEE.
[13] Imran M, Hlavacs H, Haq IU, Jan B, Khan FA, Ahmad A, Provenance based data integrity checking and verification in cloud environments. PLoS ONE 12(5): e0177576, 2017



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY