iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** II **Month of publication:** February 2018

DOI:

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 6 Issue II, February 2018- Available at www.ijraset.com

Encryption and De-Duplicating of Data in Cloud

Ms. Boomija M.D[1], Ms. Lakshmi Priya K.V[2], Ms. Suhashini Prabhu B[3], Ms. Vishali J[4]
[1] *Assistant Professor Dept. of Information Technology ,Prathyusha Engineerinng College.*
[2, 3, 4] *Students, Dept. of Information Technology, Prathyusha Engineering College.*

Abstract: Cloud computing is an online data centre for providing a large amount of computing and storage resources for various service applications with high quality. However, cloud users no longer possess their data in a local data storage infrastructure, which would result in auditing for the integrity of outsourced data being a challenging problem. To help the users complete verification of the integrity of the outsourced data has become a key issue. The secure Encryption and de-duplication techniques are used to solve this problem, from which the users can resort to a third-party auditor (TPA) to check the integrity of outsourced data and Encryption algorithm MD5 is used to secure the data outsourced. The duplication of data has been avoided using the De-duplication techniques by comparison of files in the user database using php ,this will enhance the memory efficiency of the storage and the processing time is reduced.

I. INTRODUCTION

In current days, the speeding growth of digital contents is gearing up to raise the demand for new storage and network capacities, along with an increasing need for more cost- effective and more use of storage and network bandwidth for data transfer. Now, the use of remote storage systems is gaining an expanding interest, namely the cloud storage based services, as it provides cost efficient architectures. In addition to save resources consumption in both, network bandwidth and storage capacities, many cloud services, namely Dropbox, wuala and Memopal, apply client side deduplication. This concept ignores the storage of redundant data cloud servers and reduces network bandwidth consumption associated to transmitting the same contents times.

Cloud storage service providers perform deduplication to save space by only storing one copy of each file uploaded. Should clients conventionally encrypt their files, however, savings are lost. Message-locked encryption resolves this tension[1]. However it is inherently subject to brute-force attacks that can recover files falling into a hacker hands. But customers may want their data encrypted, for reasons ranging from personal privacy to corporate policy to legal regulations[5][4]. A client could encrypt their file before storing it. But common encryption modes are randomized, making deduplication impossible since the Storage Service effectively always sees different cipher texts regardless of the data. If a client's encryption is deterministic (so that the same file will always map to the same cipher text) deduplication is possible. Cross-user deduplication, which allows more storage savings, is not possible because encryptions of different clients, being under different keys, are usually different. Sharing a single key across a group of users makes the system brittle in the face of client compromise.

II. EXISTING SYSTEM

In the traditional architecture there existed only the server and the client. In most cases the server was only a data base server that can only offer data. Therefore majority of the business logic i.e., validations etc. had to be placed on the clients system. This makes maintenance expensive. This also means that every client has to be trained as to how to use the application and even the security in the communication has to be considered. Since the actual processing of the data takes place on the remote client the data has to be transported over the network, which requires a secured format of the transfer method[3]. But transactions are considered to be "un-trusted" in terms of security, i.e. they are easy to be hacked. And also we have to consider the transfer the large amount of data through the network will leads to network traffic while transferring. In the same way sensitive data transfer is to be carried out even if there is lack of an alternative. We know that deduplication reduces the storage space at cloud server side.Data integrity is very critical issue in storage systems, because only a single instance of file has been accessed by multiple user. To ensure private information the secret sharing scheme is utilized effectively.

A. Disadvantages

The major disadvantage of existing system is the storage; it needs more memory space which is costly to maintain. The existing system depends upon sharing of data through network bandwidth. This is time consuming and expensive and have to rely on hard-drive and still not finding appropriate data to process. The data integrity and authentication is mainly important in cloud storage.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue II, February 2018- Available at www.ijraset.com

III. PROPOSED SYSTEM

We propose this web application for better security and integrity of data stored in the cloud. Mainly this web application deals with the important factor like Encryption, De-duplication, Security and Integrity of data. The sharing of data is easy but the one thing we should take care of is the security because we don't want hacker should view our data in the cloud without permission of the primary user. Here using Cloud storage, it will help the users to store their data on a
online network and can retrieve it easily from their when it is needed . Also the reason for making this web application is that, the data in cloud is not fully trustworthy and raise security issues[7]. The high cost of data storage devices and the use of data rapidly increases cloud storage .The highly secure encryption technique and very efficient De-duplication method has been used in this project for better storage of data in cloud.

IV. DESIGN AND ARCHITECTURE

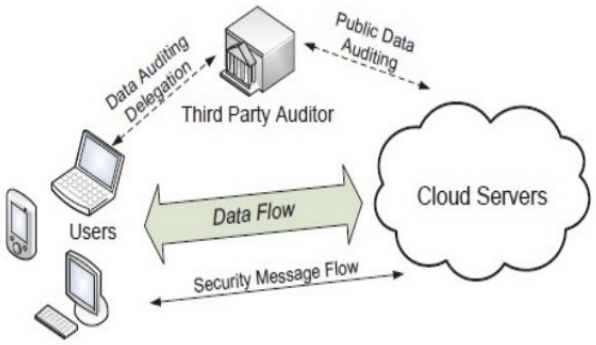


Fig. 1: The architecture of cloud data storage service

A. Users

Users or clients are who want in need of large storage of data so, they choose the cloud storage as the choice of storage for huge amount of data in online .The users create an account in cloud server provider and outsource the data into the cloud.

B. Thirty party Auditor

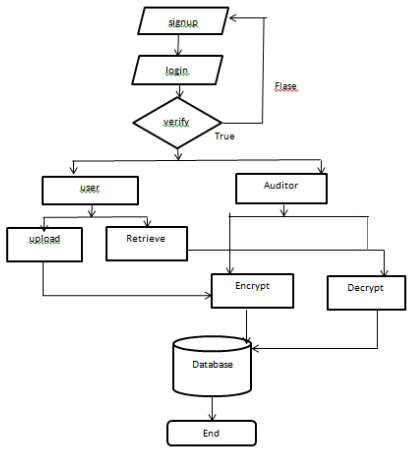
The TPA performs the main role of data integrity check. It performs activities like generating hash value for encrypted blocks received from cloud server, concatenating them and generates signature on it. It later compares both the signatures to verify whether the data stored on cloud is tampered or not. It verifies the integrity of data on demand of the users. The cloud server is used only to save the encrypted blocks of data. This proposed auditing scheme make use of MD5 algorithm for encryption and the De-Duplication is done to reduce repetition of the same file storage and increase the memory efficiency.

C. Cloud Servers

A cloud server is primarily an Infrastructure as a Service (IaaS) based cloud service model[6]. There are two types of cloud server: logical and physical. A cloud server is considered to be logical when it is delivered through server virtualization. In this delivery model, the physical server is logically distributed into two or more logical servers, each of which has a separate OS, user interface and apps, although they share physical components from the physical server. Whereas the physical cloud server is also accessed through the Internet remotely, it isn't shared or distributed.

D. Dataflow Diagram

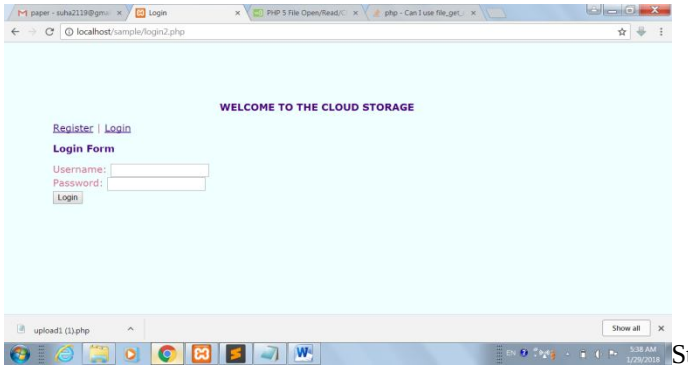
The data flow indicates the steps involved in the data storage to cloud in secure way. The user need to create an account in the cloud service provider. Later they have to choose whether to upload or retrieve the data from cloud.



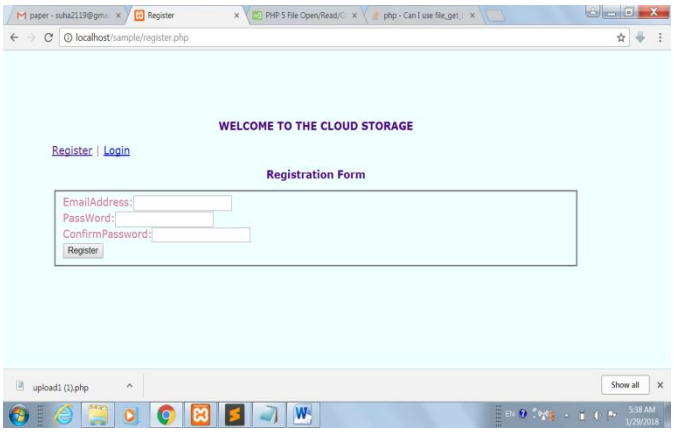
E. Implementation

The user or client first need to create an account in the cloud service provider based on the client request the space for storage will be allocated in the cloud.

Step1

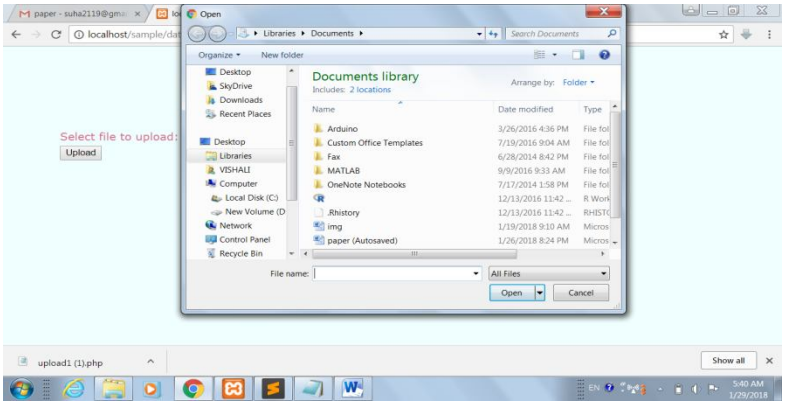


Step2:



F. Upload or Retrieve Data

Select the process you need to be performed in the cloud whether to store the data in cloud or to download the data from cloud storage.



The file will be Encrypted using the MD5 algorithm and successfully stored in the database

G. De-duplication

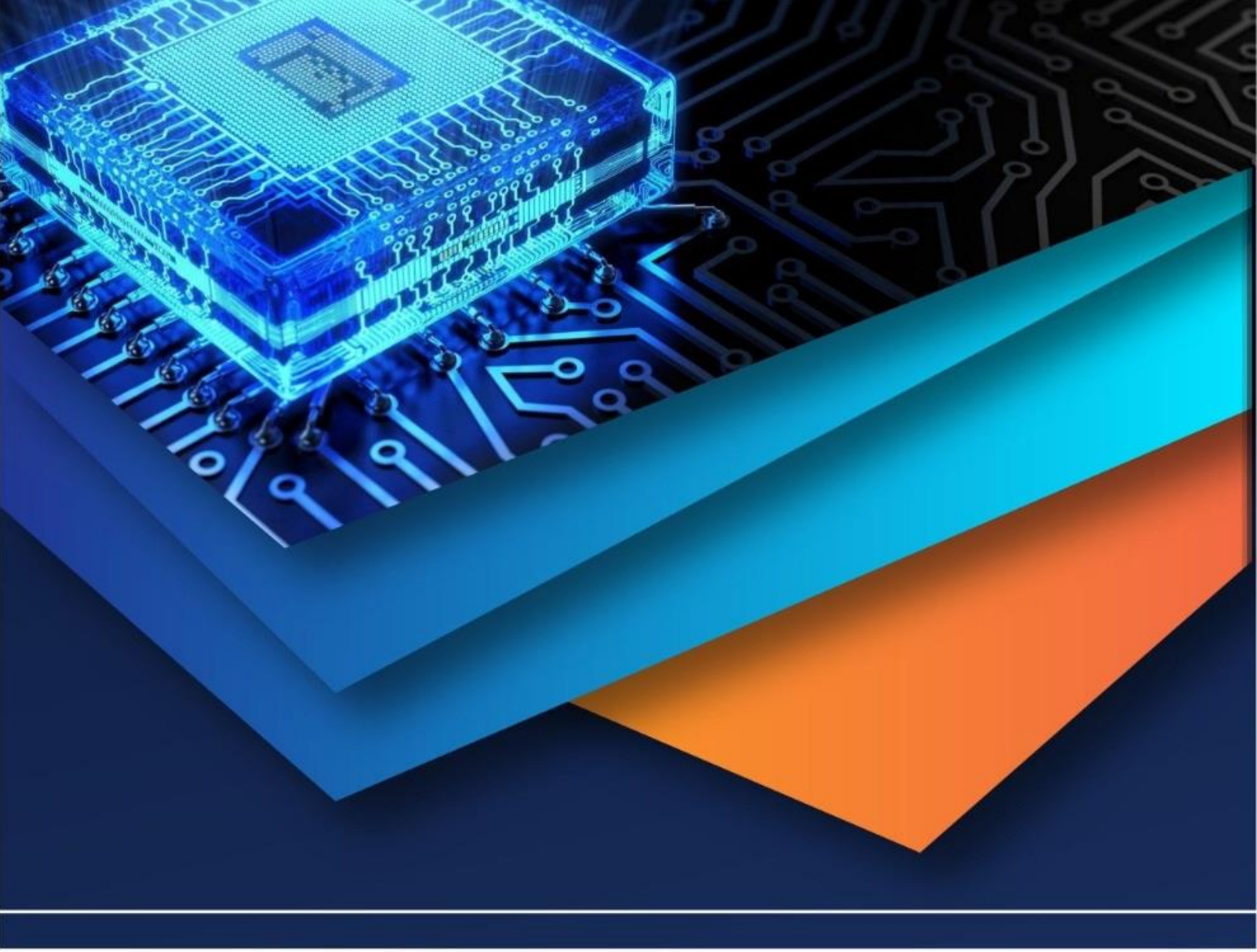
The duplication of data has been avoided by comparing the files stored in the database.

V. CONCLUSION

The aim of this application is to help the users to store their data on the cloud with confidentiality and security. De-duplication of data and encryption is the main focus in the entire application. Providing storage of large data on a cloud with multiple file sharing. Auditing helps the user to check the integrity of the data.

REFFERENCES

[1] Xuefen Liu,WenhaiSun,WenjingLou,Qingqi Pei,Yuqing Zhang,Message-locked integrity auditing on encrypted cloud deduplication storage IEEE 2017
[2] Yulin Wu,Zoe L. Jiang,Xuan Wang,S.M. Yiu,Peng Zhang, Dynamic Data Operations with Deduplication in Privac Preserving Public Auditing for Secure Cloud Storage IEEE 201
[3] Ahila.R, Sivakumari. Scloud computing security issues and techniques IJARCS 201
[4] M.Ambrust,A.Fox, R.Griffith,A.D.Joseph,R.Katz,A.Konwinski,G.lee,D.patterson, A.Rabkin,I.Stoica,and M.Zaharia,"A view of cloud computing,communication of the ACM 201
[5] J.Yuan and S.Yu,"Secure and constant cost public cloud storage auditing with deduplication"in IEEE conference on communication and network security 201
[6] S. Keelveedhi, M. Bellare, and T. Ristenpart, "Dupless: Serveraided encryption for deduplicate storage," in Proceedings of the 22Nd USENIX Conference on Security, ser. SEC'13. Washington, D.C.: USENIX Association, 2013, pp. 179–194. [Online].
[7] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in Proceedings of the 14th ACM Conference on Computer and Communications Security, ser. CCS '07. New York, NY, USA: ACM, 2007, pp. 598– 609.
[8] G. Ateniese, R. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner, Z. Peterson, and D. Song, "Remote data checking using provable data possession," ACM Trans. Inf. Syst. Secur., vol. 14, no. 1, pp. 12:1–12:34, 2011
[9] G. Ateniese, R. Di Pietro, L. V. Mancini, and G. Tsudik, "Scalable and efficient provable data possession," in Proceedings of the 4th International Conference on Security and Privacy in Communication Netowrks, ser. SecureComm '08. New York, N
[10] E. Mykletun, M. Narasimha, and G. Tsudik, "Authentication and integrity in outsourced databases," Trans. Storage, vol. 2, no. 2, pp. 107-138, 2006.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY