

# ENHANCED PROTECTION DE-DUPLICATE DATA STORAGE SCHEME FOR MULTIUSER ENVIRONMENTS IN CLOUD

M. Lawanya Shri<sup>1</sup>, Pavan Srinivas<sup>2</sup>, Priya G<sup>3</sup>

<sup>123</sup>VIT University, Vellore

## ABSTRACT

*Cloud computing is the technology which is mainly used in the industry in the data storages. This paper presents the enhanced deduplication data protection scheme in cloud storage. Deduplication which means avoiding redundant in the Cloud Storage the problem is in Multiuser Environments in this Project proposed work is how to access the Files in the Multi user Environments without uploading the file. we can access the data at any time anywhere but we need the network connection deduplication data Storage for multiuser Environments which mainly focus on the data storage and also helps to avoid not wanted data in the cloud server which mean redundant data while performing these task one more major task providing Security of the data which means the data remains same without having any modification. We here present the security at the same time data integrity using the Homomorphic Authenticated Tree for protection of cloud storage.*

## I. INTRODUCTION

The cloud data storage is broadly classified in two varieties which are namely public cloud and as private cloud public cloud which can be access by any person it is free of cost but to some extent examples of public cloud data storage mainly as the google drive, amazon and many more are there we can store the data in the public data storage we can use the storage as per the requirements because in some remote places we cannot having the resources. even though that conditions we can access the data by using the cloud data storage the main thing we need the network connectivity it is used many applications like banking e-commerce ,and other commercially sites which would provide better service to the customer and also they are improving their business. These will provide high end system which are typically related to the servers to store the data we can access any ever even in the remote places also we can easily access the data through the server .they are many more public cloud service providers which are giving service free of cost up to some extended later they would be charge. This is the technique which helps to promote their business and also given awareness to the technology people .and how to utilize the technology.it would be helpful for technology development access the data at any point of time, these is going to the future trend in the market.

De-duplicate data storage which mainly focus on eliminating the same data in the cloud storage this project main aim is to provide security of the data as well as maintaining the integrity of the data and eliminating the duplicate storage of the data this project mainly implemented using the domain of cloud data storage the data is

storage in Drive hq it is free cloud source provider and Front end is java programming Language and for data base taking the support of the SQL query and MySQL, the working area is cloud Storage the main advantage of the cloud we can access the data anywhere but we need network connection

This is develop by using the neat beans software by using the java programming language and using cloud technology the data is stored in cloud server the main focus is avoid repeated data because in current days Storage becomes more important as well in industry and other sectors.

The main advantage is of this avoid unnecessary save the bandwidth and less cost and we are able to improve the high accessibility of the files and other which is more important and also Data integrity will be maintained the data would not be altered and the project mainly works on the user and subsequent user need to upload the same context which is not allowed it will indicates already the context is already present in the server so the other send request to the actual user he is having ultimate authority whether to access the data or not it .used on the cloud computing domain which has mainly Infrastructure as service, the organization need not be maintain the data which means physical infrastructure instead of the cloud service provide will maintain all the settings of the user pay as per their usage by this mechanism it is advantage of the business person he is not able to set up any infrastructure and also unnecessary burden everything will be concern the cloud provider and also they reduces the employees cost for making employees it is also avoid the risk and helps to achieve the best optimal results .

Proxy login is the module which is emerged in this project the which help to gives access to the subsequent user to access the data owner file.it mainly re encryption of the application which the other user is having access to use the file of the actual data owner the proxy login is the module it helps to authority to the subsequent user access the files with proper verification .by doing these we can eliminate the de-duplicate piece of data and improves efficiency and performance of the system would be in terms of response time high response time and also utilized in all the applications in the cloud domain in the services and also necessary business needs of the organization.

## **II. PROPOSED SYSTEM**

The Proposed system was implemented in Cloud domain the main intention is access the files in multiuser Environments we are implements the system the user upload the data in the server which is in the form of file the other user has to want to upload the same file it is not excepted which is already exist .by access the subsequent user these user need to take permission from these data owner the data owner give permission after verifying the he is actual user or not the subsequent user send request to the data owner after expected the requested the other user is eligible to access the file while performing these process security and data integrity is the concern in this proposed system by maintaining all these factors we are using the model which is named Enhanced protected deduplication data for multiuser Environments using macs codes for providing better security and also the data is split in two blocks by using the HAT , for these we achieve the avoiding the wastage of bandwidth. Improving performances, efficiency and also help to better experiences for the user end in the terms of the response time and more things the data waste would be reduced. The system design is needed

mainly because to know the functionality of the system before developing any modifications required by the client side we can easily modify .after started of the project they need changes it become difficult because enter pattern would be changed .so before developing any product the design team analyse the specifications and requirements they release a blue print the customer will see this blue print any modifications wants to take place through these, they discuss after starting the things they said it is so what costly mechanism to develop.

### **III. PRE-PROCESS**

This is the phase on the primary starts on the system the user need to register their details in the application after that the cloud admin will validate their user specification, whether the user is correct or not and is the user is valid then it gives access in cloud storage, the cloud server will decides whether to upload the file or not handle stage, clients expect to transfer their nearby records. The cloud server chooses whether these records ought to be transferred. On the off chance that the transfer procedure is truly, go into the transfer stage; generally, go into the deduplication stage. The pre-handle stage, a client runs the instatement calculation. On the off chance that the document does not exist, the client goes into the transfer stage. Something else, the client goes into the deduplication stage. The document declared by a client in the pre-handled stage occur in the cloud server, the client enters into the deduplication stage and goes to the deduplication

### **IV. USER UPLOAD TO CLOUD**

The transfer stage, the records to be transferred it would not be occur in the cloud server. The first client encodes the nearby documents and transfers them to the cloud server. These stages end ordinarily, i.e., clients get done with transferring in the transfer stage, or they pass the confirmation in the deduplication phase is the transfer stage, the client transfers C and T to the cloud server and just stores e locally. Take note of that, e is a component of little consistent size and can be encoded and put away in the cloud server. Conversely with that requires clients having or downloading a structure which has logarithmic size of the quantity of document hinders, all proprietors of the record can run the deduplication convention, the checking convention. Firstly checks whether these document is available in the server if it there in the same context it notifies the already exist in the application so no need to upload just take permission form the subsequent user is enough.

### **V. DEDUPLICATION**

The deduplication stage, the documents to be transferred as of now involved in the cloud server. The consequent clients have the documents locally and the cloud server stores the confirmed structured of the records. Consequent clients necessity to persuade the cloud server which they claim the records with any transferring them to the cloud server. we outline a enhanced validated structure called Homomorphic Authenticated Tree (HAT), to decrease the correspondence cost in both the confirmation of capacity stage and the deduplication stage with comparable calculation cost. The necessity of cross-client deduplication on the customer side is identified. This shows clients can avoid the transferring procedure and get the responsibility for promptly, the

lengths of the transferred records as of now occur in the cloud server. This strategy can lessen storage room for the cloud server, and spare transmissions transfer speed for clients

## **VI. PROOF OF STORAGE**

The verification of capacity stage, clients just have a little steady size metadata local and they need to check whether the records are dependably put away in the cloud server without any burden to downloading them. The records may not be transferred by these clients, but rather they pass the deduplication stage and demonstrate that they have to involve the proprietorships of the files .the refresh stage and to be the confirmation of capacity stage could be executed numerous circumstances in the span cycle of a document. Hence the proprietorship is confirmed, the clients can discretionarily enter the refresh stage and the evidence of capacity stage without keeping the first documents locally demonstrated its security in the irregular prophet display. The hypothetical and test comes about demonstrate that our KeyPoS usage is productive, particularly when the document estimate and the quantity of the tested pieces.

## **VII. PROXY RE-ENCRYPTION**

This modules which mainly deals with the re-encryption process basically when the user upload their documents in the form of the files these actual user can be called as the data owner the other authorized user wanted to upload the same context in the form of file the user send request to the proxy .here proxy acts as medium between the data owner and subsequent user. the subsequent user send request to the proxy .once receives the request from the user the proxy will evaluate here there is an actual user or not after confirm the details the proxy will accept their user request then the subsequent user gets ownership of the data .now he is having authority to download the context .by these module we can reduce the storage space instead of uploading multiple files the other user simply skip the files and gets the ownership authority of the data on the context. These proxy login is used mainly to the access to the user based on the request of the file .when the other the data owner want to access the file it means the subsequent user firstly the other user send request to the proxy .when the request received by proxy .in these module it will Be activate the authority to access the files .once in this module it activated .the subsequent user is able to access the context which means the other user got the ownership of the file at that point of time the user is able to view the file and edit thee file as the delete the context similarly able to download the aim working of these module is provides the data ownership of the other users which are access through the systemthe user pay as per their usage by this mechanism it is advantage of the business person he is not able to set up any infrastructure and also unnecessary burden everything will be concern the cloud provider and also they reduces the employees cost for making employees it is also avoid the risk and helps to achieve the best optimal results

Proxy login is the module which is emerged in this project the which help to gives access to the subsequent user to access the data owner file.it mainly re encryption of the application which the other user is having access to

use the file of the actual data owner the proxy login is the module it helps to authority to the subsequent user access the files with proper verification .by doing these we can eliminate the de-duplicate piece of data and improves efficiency and performance of the system would be in terms of response time high response time and also utilized in all the applications in the cloud domain in the services and also necessary business needs of the organization.

### **VIII. CONCLUSION AND FUTURE WORK**

Outline of an enhanced system for capacity in Cloud is deduplication procedure. Deduplication helps in sparing the storage space. This application helps in simple upkeep of information on the cloud stage so that no copy documents are spared in the Cloud. With the development of Cloud processing, stockpiling assets of product machines can be proficiently used. This permits each association to manufacture its own private cloud for an assortment of purposes. With a specific end goal to better use the constrained stockpiling accessible in a private cloud, an appropriate approach for advancement must be utilized. Deduplication is actualized on the storage room of Amazon Cloud. In this arrangement, deduplication is conceivable at client's container level. The outcomes propose an impressive reserve funds in the storage room of Cloud is accomplished. Deduplications of records are accomplished for every client basin. Future upgrade is accomplishing deduplication on the entire distributed storage. Currently, enhanced method for capacity has been tried just for content records. In future, it can be additionally reached out to bolster records of different sorts.

### **REFERENCES**

- [1] G. Ateniee, R. Di Pietro Efficient Provable Data Possession," in Proc. of SecureComm, pp. 1– 10, 2008.
- [2] G. Ateniee, S. Kamara, and J. Katz, "Confirmation of capacity from homomorphic recognizables proof conventionals," in Proc. of ACRYPT, pp. 319–333, 2009.
- [3] C. Erway, A. Ku" pcu and R. Tamassia, "Dynamic provable information ownership," in Proc. of CCS, pp. 213–222, 2009.
- [4] R. Tamassia, "Validated Data Structures," in Proc. of ESA, pp. 2–5, 2003.
- [5] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Empowering open certainty and information progression for capacity security in distributed computing," in Proc. of ESORICS, pp. 355–370, 2009.
- [6] F. Armknecht, J.- M. Bohli, G. O. Karame, Z. Liu, and C. A. Reuter, "Outsourced confirmations of retrievability," in Proc. of CCS, pp. 831–843, 2014.
- [7] H. Shacham and B. Waters, "Reduced Proofs of Retrievability,"
- [8] Z. Mo, Y. Zhou, and S. Chen, "A dynamic confirmation of retrievability (PoR) conspire with  $o(\log n)$  many-sided quality," in Proc. of ICC, pp. 912– 916, 2012.
- [9] E. Shi, E. Stefaonov, and C. Papaamanthou, "Commons sense elementination verifications of retrievability," in Proc. of CCS, pp. 325–336, 2013.