



# Reducing the Duplication Data for Secure Authorization in Hybrid Cloud

<sup>1</sup> J. Jinu Sophia, <sup>2</sup> R. Angelin Hannah, <sup>3</sup> J. Daisy, <sup>4</sup> K. Jebima Jessy

<sup>1</sup> Assistant Professor, Dept. of IT, Kings Engineering College,

<sup>2,3,4</sup> UG Scholars, Dept. of IT, Kings Engineering College.

<sup>1</sup>[Jinu42@gmail.com](mailto:Jinu42@gmail.com), <sup>2</sup>[hannah.joe195@gmail.com](mailto:hannah.joe195@gmail.com), <sup>3</sup>[daisyjohn@gmail.com](mailto:daisyjohn@gmail.com), <sup>4</sup>[jessyjebima@gmail.com](mailto:jessyjebima@gmail.com)

**ABSTRACT:** Deduplication of data's is a technique used for eliminating duplicate copies and to decrease the space in storing memory. It is also used in saving bandwidth. The problem in the paper is that the data which is authorized or not. The solution of the data is proposed to the authorized by the methods such as authorized duplicate check and differential authorization. In these methods the private key is sent by the admin mail request to the user's mail after the file is shared. The authorized duplicate check is used for new deduplication construction in the hybrid cloud architecture various tested experiment are used as a prototype. The data security is used as protection for deduplication. The authorization of data should be protected by the authorized duplication check scheme.

## I. INTRODUCTION

Cloud computing, also an demand computing is a kind of internet based computing that provides shared processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitin, on demand access to a shared pool of configurable computing resource cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in third party data centres.

Cloud computing is a delivery of hosted service over the internet. The practice of using a network of remote service hosted on the internet do store, manage and process data, rather than a local server or a personal computer is mentioned in Fig1. Main benefits of cloud computing are self service provisioning end user can spin up computing resources for almost any type of workload on demand, elasticity, pay per use. These are three types of cloud computing they are

1. Private cloud
2. Public cloud
3. Hybrid cloud

A private cloud services destructed from a business data centre to internal use. Private cloud is also known as internal cloud or corporate cloud.

Public cloud is a deliver services to multiple organization, a private cloud is dedicated to a single organization. Public cloud is a third party provides delivers the cloud service over the internet. A hybrid cloud is the combination of private and public cloud. It is to create a unified, automated, scalable environment which takes advantage of all that a public cloud infrastructure can provide which still maintaining control over mission critical data.



Fig1: Welcome page of the website

## II. STUDY METHODOLOGY

### A. User authorization process

In the authentication process, the user should sign up in the website to create an account by new user which is mentioned in Fig2. Now, the user is a member of the website. The user can view the website and this specified user can upload the file to the website as mentioned in Fig3. The user can share to other users as he is authenticated. The authentication process consists of two parts: an interactive logon process and a network authentication process. A successful user authentication depends on both of these processes. The interactive logon process involves user identification to either a domain account or a local computer network authentication confirms the user identification to any network service that the user is attempting to access to provide this type of authentication. Many mechanisms support including Kerberos, secure socket layer / transport layer security (SSL/TLS) and for compatibility with Windows NT 4.0, LAN Manager users who use a domain account do not see network authentication. Users who use a local computer account must provide credentials (such as username and password) every time they access a network

resource.

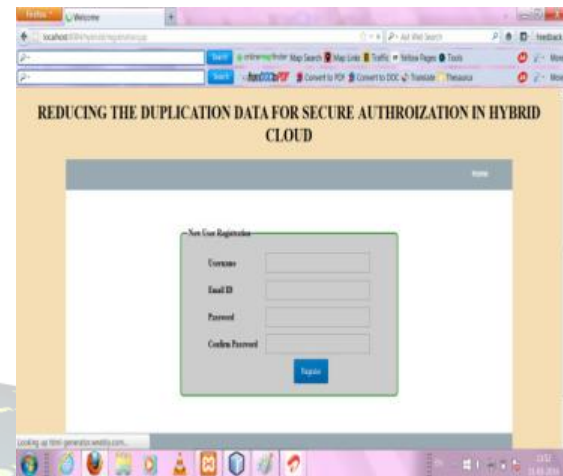


Fig2: User registration process

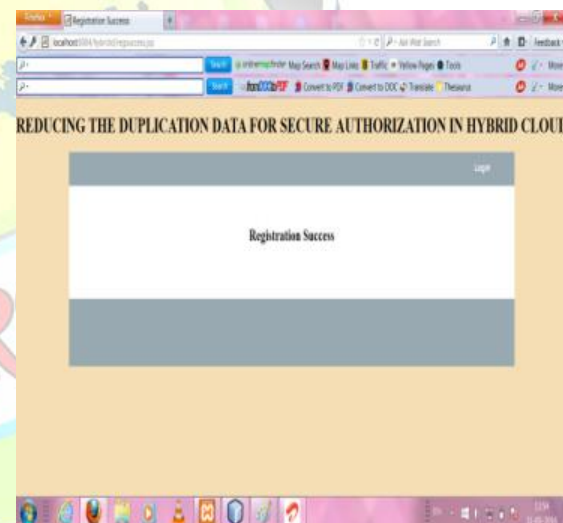


Fig3: Registration of the user is successful.

By using the domain account, the user credentials that are automatically used for single sign on. A data user refers to a person involved in accessing and investigating integrated data set for statistical and research purpose. Data users are directly involved in analysing integrating datasets at the unit record level to conduct and undertake research. Data users have a number of rights and responsibility associated with accessing and using integrated datasets that involved use of at least one dataset. It is the right of data users to consult with data custodians and the integrating



authority or any material changes or updates to a data integration project. Data user must be aware of and understand sanction which apply for the attends to identify individual or organisation.

### B .Uploading of data

The authenticated user can able to upload the data or files to the website as mentioned in Fig4. The data can be of text video and audio or any other file the member of the website can only be able to sign into its account to upload the files.Any text files document can be uploaded by the given data size and storage space the uploading of data is three stage process as mentioned in Fig5. (1) The data file must first be screened for errors .if the report produce no users and the user accept all warnings reported in the files. The file can pass to the next stage. (2) The data which is plotted allow the user to visually check for outliers.If the user accepts the data and any outliers, the file can pass to next stage(3) the file is now ready for uploading the text document the user will be prompted to conform on screen.In computer networks to upload is to send data to a remote system such as server or another client so that the remote system can store a copy.



Fig 4: File Uploading Process



Fig5: Uploading of file is successful.

To upload something is generally understood to mean “putting something on the internet “for example,uploading photographs Facebook. The uploading is use by some online file hosting services. It is also used when the local computer has a slow connection to the remote system but they have a fast connection between them. After you have created and registered your data feed,we can upload a data file into it.

### C.Sharing of data

The authorized user can able to share the data's only to the authorized user of that website .The sharing of data's can be processed only if the authorized user had already uploaded the data is mentioned in Fig6.The authorized users share the data by mentioning the id of the other user.Data sharing is the ability to share the same data resource with multiple applications or users. It implies the data are stored in one or more servers in the network and that there is some software locking mechanisms that prevents the same set of data from being changed by two people at the same time.Data sharing is a primary feature of a database management system.In this sharing of data,the data's which is shared is protected by a key which is provided to the authorized user by the user who shared the file or document. The key is provided through the mail which is used by the authenticated use. If the key is not correct the file cannot be



downloaded by the user. This is the process of security in the sharing of data.



Fig6:File sharing Process.



Fig7: Sharing of file is successful.

### D.Existing file system

If the user has the file in their account either it is uploaded or downloaded the same document or file cannot be existed in that account. It delivers the message as that the document is already present and the old file exists. This will provide awareness to the user and also it eliminates the repeating copies of data which is the process of data deduplication is shown in the Fig8.



Fig8: Detection of Deduplication.

### III.PROPOSED METHODOLOGY

Every data can be easily uploaded and they can be shared by the authorized users but the files can only be downloaded by the authenticated user who is having the private key of that particular document. It provides the security or protection to the files which have been shared to the other users. And also they existing file cannot be uploaded as it creates and awareness message to the users is shown in the below Fig9.

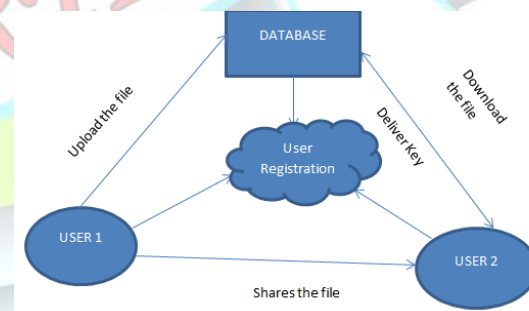


Fig9:System Architecture.

### IV. RELATED WORKS

In disk based backup data deduplication promises to reduce the transfer and storage of redundant data, which optimizes network bandwidth and storage capacity. Storing data more efficiently disk lets you



retain data for longer periods or “recapture” data to protect more application with disk based backup, increasing the likelihood the data can be recovered rapidly. Transferring less data over the network also improves performance. For example, data protection processes, data deduplication may create a unique data object with the hash algorithm and check that fingerprint against a master index. Unique data is written to storage and only a pointer to the previously written data is stored. In data duplication product trends in data backup Stephen Forsett, director of data practice at storage consultancy contour said deduplication will become a necessary technology for backup and take big steps into archiving in 2009. But dedupe isn't yet ready for many types of primary storage to the chagrin of some of the storage customers that he talks to. This technology is still developing and will always be developing. Introduction to data deduplication in windows server 2012. Data deduplication that enables to efficiently store, transfer and backup use data. It is the result of an extensive collaboration with Microsoft research and after two years of development and testing we now have state of the art deduplication that uses variable chunking and compression and it can be applied to your primary data. How to avoid data deduplication fiasco. If a storage administrator doesn't have a solid understanding of data deduplication there's a good chance the deployment of this technology will be wasted one. Coupled with virtual tape library (VTL) technology, deduplication was once willed as the solution to all backup spatial needs. Unfortunately, the market for deduplication hasn't take off as expected mainly due to two things

- (1) How the benefits have been communicated.
- (2) The basic understanding of the technology itself

## **V. CONCLUSION AND FUTURE ENHANCEMENT**

This paper has the set up to prevent the data deduplication, the key establishment for security purpose and eliminating the repeated copies of data to save the bandwidth. The client and the server manages the database providing the prevention of data deduplication. Future enhancement can be made in this area by avoiding the deduplication while the original file occurs.

## **REFERENCES**

- [1]. S. Bugiel, S. Nurnberger, A. Sadeghi, and T. Schneider, “TwinClouds: An architecture for secure cloud computing,” in Proc. Workshop Cryptography Security Clouds, 2011, pp. 32–44.
- [2]. Christo Ananth, A. Nasrin Banu, M. Manju, S. Nilofer, S. Mageshwari, A. Peratchi Selvi, “Efficient Energy Management Routing in WSN”, International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE), Volume 1, Issue 1, August 2015, pp:16-19
- [3]. 8jiiK. Zhang, X. Zhou, Y. Chen, X. Wang, and Y. Ruan, “Sedic: Privacy-aware data intensive computing on hybrid clouds,” in Proc. 18th ACM Conf. Comput. Commun. Security, 2011, pp. 515–526.
- [4]. J. Yuan and S. Yu, “Secure and constant cost public cloud storage auditing with deduplication,” IACR Cryptology ePrint Archive, 2013:149, 2013