

**A SIGNIFICANT REDUCTION OF CLOUD STORAGE BY  
ELIMINATION OF REPETITIVE DATA****M.Rajashekar Reddy<sup>1</sup>, B.Ramya<sup>2</sup>**<sup>1</sup>M.Tech Student, Dept of CSE, Arjun College of Technology & Sciences, Hyderabad, T.S, India<sup>2</sup>Assistant Professor, Dept of CSE, Arjun College of Technology & Sciences, Hyderabad, T.S, India**ABSTRACT:**

Because of cloud system developing into an established technology, an increase amount of data is stored within cloud system and is shared by users by particular privileges specifying access rights of stored information. The process of de-duplication symbolizes a determined data compression for taking away of duplicate copies of repetitive data within storage system. We aim at working out deduplication problem by means of differential advantages in cloud system and hence consider a cloud design which is hybrid structure consisting of public along with a private cloud. An approved data deduplication was initiated in our work to defend data security by means of inclusion of differential opportunities of users within duplicate check. Contrasting from traditional works, private cloud is considered as a proxy for data owner to execute duplicate check by differential privileges and such a design is useful and has offered much attention.

***Keywords: Cloud system, Private cloud, Deduplication, Hybrid structure, Duplicate check, Differential privileges.***

## 1. INTRODUCTION:

Rather than maintaining of several data with similar content, deduplication concept removes redundant data by means of managing one physical copy and refers other redundant data towards that copy. The knowledge of data deduplication helps in improving of storage utilization and moreover it is apt towards network data transfers to decrease number of bytes that have to be sent [1]. In the earlier systems of deduplication, differential authorization duplicate check, was not supported which is the significant factor to be considered in most of the applications. And in such systems, each user is provided by a set of privileges throughout the period of system initialization. Earlier systems of deduplication on the basis of convergent encryption, even though provides privacy to some level, however do not manage duplicate check by differential privileges. Deduplication technique on the basis of convergent encryption technique, there are no differential privileges considered. In our work we intend at working out deduplication problem by means of differential advantages in cloud system and hence consider a cloud design which is hybrid structure consisting of public along with a private cloud. For

defending sensitive data confidentiality while managing the process of deduplication, technique of convergent encryption was introduced for the purpose of encryption of data earlier than outsourcing. Proposed system consists of twin clouds such as public cloud as well as private cloud. For securing of data protection, our work makes an initial effort to formally for addressing deduplication difficulty of approved data.

## 2. METHODOLOGY:

The knowledge of data deduplication method represents a focused data compression intended for elimination of duplicate copies of repetitive data within storage system. deduplication method occurs at moreover file level or else block level. While the technique of data deduplication holds several benefits, however the issued regarding security as well as privacy happens since users' data are prone to various attacks [2][3]. Traditional techniques of encryption, while offering confidentiality of data, are not compatible with data deduplication method. Traditional techniques of encryption necessitate several users to encrypt their information by personal keys and so data copies of

numerous users that are similar will guide to a variety of cipher-texts, making of deduplication impractical. Technique of convergent encryption was introduced to put into effect data privacy while making possibility of deduplication. Deduplication procedure on the basis of convergent encryption system, there are no differential privileges considered. Technique of convergent encryption allows cloud to carry out deduplication on ciphertexts as well as proof of ownership for preventing the unapproved user towards accessing of file. Deduplication techniques of earlier works based on convergent encryption, although providing privacy to some level; however do not manage duplicate check by differential privileges. Here in our work we intend to solve the deduplication problem by means of differential advantages in cloud system and hence consider a cloud design which is hybrid structure consisting of public along with a private cloud. In our proposed system, contrasting from the earlier works, private cloud is considered as a proxy for data owner to execute duplicate check by differential privileges and such a design is useful and has offered much attention. The owners of data outsource their data storage by means of employing public cloud while

data procedure is managed within private cloud. The innovative and efficient deduplication system that supports differential duplicate check was introduced in the hybrid cloud structural where the provider of Storage-cloud service resides within public cloud. Here user only carries out duplicate check for files that are marked by corresponding privileges.

### **3. AN OVERVIEW OF PROPOSED SYSTEM:**

In the modern times, providers of cloud service usually provide storage that is mostly accessible and the resources of computing at an extremely low price. One of the important issues to be considered regarding cloud storage services is managing of increased data volume. For managing of scalable data within cloud system, de-duplication has proved to be an effective means that has gained more attention in recent times. The process of de-duplication occurs at block level that removes duplicate data blocks occurring in non-identical files [5]. This method represents a determined data compression intended for elimination of duplicate copies of repetitive data within storage system. deduplication method. In our work

approved data deduplication was introduced to defend data security by means of inclusion of differential opportunities of users within duplicate check. We intend at working out deduplication problem by means of differential advantages in cloud system and hence consider a cloud design which is hybrid structure consisting of public along with a private cloud. In our system private cloud is considered as a proxy for data owner to execute duplicate check by differential privileges and such a design is useful and has offered much attention. For securing of data, our work address deduplication difficulty of approved data. The deduplication system supporting differential duplicate check was set up within hybrid cloud structural where the provider of Storage-cloud service resides within public cloud and in this only user carries out duplicate check for files that are marked by corresponding privileges. In the security representation of our systems private cloud is supposed to be honest however curious. Proposed structure for effective Deduplication is a new structural design for data deduplication within cloud computing system that consisting of twin clouds such as public cloud as well as private cloud. In the proposed structure

deduplication technique is normally employed for data backup as well as applications for disaster recovery whereas reducing storage space to a great extent. These systems are prevalent and more appropriate towards backing up of user file as well as synchronization applications than comfortable storage abstractions. There are different entities in proposed system such as users, private cloud as well as provider of storage-cloud service provider within public cloud as in fig1. Users contain permission towards private cloud server, which is a semi trustworthy third party supporting in performing of deduplicable encryption by means of generating file tokens for the requested users [4]. To upload a file, user act upon file-level duplicate check and when file is duplicate, subsequently the entire blocks have to be duplicates if not, user performs block-level duplicate check and make out exceptional blocks to be uploaded. Provider of storage-cloud service provides services of data outsourcing and stores up data in support of users. In our work we imagine that the entire files are imagined to be protected against public and private cloud since the files are sensitive [6].

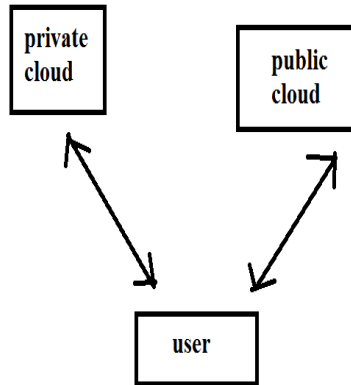


Fig1: Design of proposed system.

#### 4. CONCLUSION:

For supporting scalable data within cloud structure, de-duplication has proved to be an effective means that has gained more attention in recent times. This technique represents a determined data compression intended for elimination of duplicate copies of repetitive data within storage system and occurs at moreover file level or else block level. In our work we propose deduplication problem by differential advantages in cloud system and hence consider a cloud design which is hybrid structure consisting of public along with a private cloud. For securing of data privacy, our work makes a try to formally for addressing deduplication difficulty of approved data. The innovative deduplication system that supports differential duplicate check was introduced in hybrid cloud structural where the provider

of Storage-cloud service resides within public cloud. Here user carries out duplicate check for files that are marked by equivalent privileges. In the proposed structure deduplication technique is normally employed for data backup as well as applications for disaster recovery whereas reducing storage space to a great extent.

#### REFERENCES

- [1] M. Bellare and A. Palacio. Gq and schnorr identification schemes: Proofs of security against impersonation under active and concurrent attacks. In CRYPTO, pages 162–177, 2002.
- [2] S. Bugiel, S. Nurnberger, A. Sadeghi, and T. Schneider. Twin clouds: An architecture for secure cloud computing. In Workshop on Cryptography and Security in Clouds (WCSC 2011), 2011.
- [3] J. R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M. Theimer. Reclaiming space from duplicate files in a serverless distributed file system. In ICDCS, pages 617–624, 2002.
- [4] R. D. Pietro and A. Sorniotti. Boosting efficiency and security in proof of ownership for deduplication. In H. Y. Youm and Y. Won, editors, ACM Symposium on Information, Computer and Communications Security, pages 81–82. ACM, 2012.
- [5] S. Quinlan and S. Dorward. Venti: a new approach to archival storage. In Proc. USENIX FAST, Jan 2002.
- [6] A. Rahumed, H. C. H. Chen, Y. Tang, P. P. C. Lee, and J. C. S. Lui. A secure cloud backup system with assured deletion and version control. In 3rd International Workshop on Security in Cloud

Computing, 2011.



**M.Rajasekhar Reddy** graduated in 2013 in B.Tech,CSE, Arjun College of Technology & Sciences,R.R. Dist.



**B.Ramya** Graduated in B.Tech CSE in 2008 from Swami Ramanand Thirde Institute of Science, NLG Dist. She received Masters Degree in M.Tech [IT] Aurora Scientific Technology & Research Academy,Hyd. Presently She is working as Associate Professor in CSE Dept. in Arjun College of Technology & Sciences, Hayathnagar,R.R. Dist Telangana State, India.