

Towards A Novel Hybrid Approach for Removing Duplicate Copies of Repeated Data

D.SRAVANI

M.Tech Student

Dept of CSE

Jagruiti Institute of Engineering and Technology
Hyderabad, T.S, India

P.GIRIDHAR

Associate Professor

Dept of CSE

Jagruiti Institute of Engineering and Technology
Hyderabad, T.S, India

Abstract: Previous systems cannot support differential authorization duplicate check, in several applications. In the recent times, structural design was offered that consisting of twin clouds for effective outsourcing of data as well as arbitrary computations towards an untrustworthy commodity cloud. With the introduction of cloud computing, efficient secure data deduplication has attracted much concentration in recent times from research community. Data deduplication is a committed data compression technique that is generally introduced for eliminating duplicate copies of repeating storage data. Distinct from established systems, private cloud is offered as a proxy towards permitting data owner to securely execute duplicate check by differential privileges and hence this architecture is useful and has attracted much consideration from researchers. In our work we solve difficulty of deduplication by differential privileges within cloud computing, we imagine a hybrid cloud structural design consisting of a public cloud and private cloud.

Keywords: Cloud computing, Public cloud, Outsourcing, twin clouds, Data storage.

I. INTRODUCTION

The technique is utilized for improving of storage utilization and can moreover be functional to network data transfers for reduction of number of bytes that have to be sent. Rather than preserving of multiple data copies with identical content, deduplication method removes redundant data by means of preserving of only one physical copy and referring of other redundant data towards that copy [1]. Conventional encryption, while provision of data confidentiality, is unsuited with data deduplication. Particularly, conventional encryption necessitates various users to encrypt their information by their own keys as a result, matching data copies of various users will direct to distinctive cipher-texts, making deduplication not possible. While encryption process is deterministic and is derived from data content, matching data copies will produce similar convergent key and for this reason the similar ciphertext. To build efficient data management in cloud computing, deduplication was regarded as a renowned procedure that has gained increased attention in recent times. Convergent encryption has been projected to implement data privacy while making deduplication practicable. Earlier deduplication systems cannot maintain differential authorization duplicate check, which is significant in numerous applications. Even though data deduplication provides various benefits, security as well as privacy concerns happen since users' sensitive data are vulnerable to insider as well as outsider attacks. In our work we aim at resourcefully solving difficulty of deduplication by differential privileges within cloud computing, we imagine a hybrid cloud design consisting of a public cloud and private

cloud. No differential privileges were considered in deduplication on the basis of convergent encryption method. In authorized deduplication system, each user is provided a set of privileges all through system initialization

II. METHODOLOGY

In the recent times, providers of cloud service recommend highly obtainable storage as well as extremely parallel computing resources at comparatively low expenses. Data deduplication is a dedicated data compression method that is mainly introduced for eliminating duplicate copies of repeating storage data. It is employed for elimination of duplicate copies of repeatative information, and was extensively used in cloud storage to decrease amount of storage space as well as to save bandwidth. Deduplication system can take place at block level, which removes duplicate blocks of data that take place in non-identical files. Conventional systems of deduplication on the basis of convergent encryption, even though providing confidentiality to a degree, do not maintain duplicate check by differential privileges [2]. No differential privileges were considered in deduplication on the basis of convergent encryption method. For the most part, conventional encryption necessitates various users to encrypt their information by their own keys as a result, matching data copies of various users will direct to distinctive cipher-texts, making deduplication not possible. In the recent times, architecture was offered that consisting of twin clouds for effective outsourcing of data as well as arbitrary computations towards an untrusted commodity cloud. Cloud computing makes a provision of

limitless virtualized resources towards users as services across Internet, while hiding platform as well as implementation details. To better defend data security, our work aim at resourcefully solving difficulty of deduplication by differential privileges within cloud computing, by imagining a hybrid cloud design consisting of a public cloud and private cloud [3]. Altered from conventional deduplication systems, differential privileges of users are additionally considered in duplicate check besides data itself. A novel deduplication structure supporting differential duplicate check is projected in hybrid cloud structural design where Storage-cloud service provider resides in public cloud. The user is only authorized to achieve the duplicate check in support of files that are marked with equivalent privileges. Unlike the existing systems of data deduplication, private cloud is concerned as a proxy to permit data owner to securely execute duplicate check by differential privileges and hence this architecture is useful and has attracted much concern from researchers. The data owners merely outsource their information storage by means of utilizing public cloud whereas data procedure is managed within private cloud.

III. AN OVERVIEW OF ROPOSED SYSTEM FOR SOLVING OF DEDUPLICATION DIFFICULTY

Deduplication scheme removes redundant data by means of preserving of only one physical copy and referring of other redundant data towards that copy to a certain extent than preserving of multiple data copies with the same content. Data deduplication is the significant data compression method for elimination of duplicate copies of repeatative information, and was extensively used in cloud storage to decrease amount of storage space as well as to save bandwidth. In our work, we believe to address authorized deduplication difficulty above data in public cloud hence we aim at resourcefully solving difficulty of deduplication by differential privileges within cloud computing, we imagine a hybrid cloud design consisting of a public cloud and private cloud. Since cloud computing technology is widespread, a rising quantity of data is being stored and shared by users in cloud with specific privileges. One key challenge regarding cloud storage services is management of growing data volume [4]. Unlike conventional deduplication systems, differential privileges of users are additionally considered in duplicate check besides data itself.

A user sends duplicate-check tokens in the direction of public cloud for approved duplicate check. In a storage system that supports deduplication, user only uploads exceptional data however does not upload any duplicate data to keep upload bandwidth, which might be owned by similar user or else different users. Private is a

recent entity introduced for facilitating user's protected usage of cloud service.

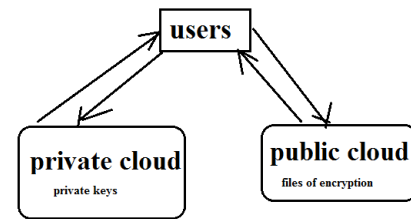


Fig1: provision of design for approved Deduplication.

The conventional encryption system requires various users to encrypt their information by their own keys as a result, matching data copies of various users will direct to distinctive cipher-texts [5]. To protect privacy of sensitive data though supporting deduplication, convergent encryption method has been projected to encrypt data earlier than outsourcing. A novel structure supporting differential duplicate check is projected in hybrid cloud structural design where Storage-cloud service provider resides in public cloud. There are three entities that are described in our system, such as users, private cloud and storage cloud service providers within public cloud as revealed in fig1. storage cloud service providers performs deduplication by means of checking if contents of two files are similar and accumulates only one of them.

IV. CONCLUSION

A design in recent times was offered that consisting of twin clouds for effective outsourcing of data as well as arbitrary computations towards an unreliable commodity cloud. We work out difficulty of deduplication by differential privileges within cloud computing, we imagine a hybrid cloud structural design consisting of a public cloud and private cloud. Distinct from established systems, private cloud is offered as a proxy towards permitting permit data owner to securely execute duplicate check by differential privileges and hence this architecture is useful and has attracted much consideration from researchers. In recent times, providers of cloud service advise greatly accessible storage as well as extremely parallel computing resources at comparatively low expenses. Even if data deduplication provides a variety of benefits, security as well as privacy concerns happen since users' sensitive data are vulnerable to insider as well as outsider attacks. Different from the traditional systems of deduplication, differential privileges of users are additionally considered in duplicate check besides data itself.

V. 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.

AUTHOR's PROFILE



Sravani .D: Pursuing M.Tech (C.S.E) in Jagruti College of Engineering and Technology