Volume 5, Issue 5, May 2015 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Implementation of Efficient Secure Storage of Deduplicated Data Technique: Using Hybrid Cloud Approach

Santoshi S Patil
Student, M.Tech (Software Engineering),
NHCE, Bangalore, India

Prof. Shwetha K S.
Assistant Professor, Information Science &
Engineering Department, NHCE, Bangalore, India

Abstract*— *Data deduplication is one of useful data compression scheme for eliminating redundancy in the storage, and has been very popular and cheap in cloud system to reduce the cloud size and save speed. For Security of sensitive data with supporting deduplication, the concept of convergent encryption technique has been introduced to encrypt the data before uploading.Although data deduplication brings a lot of benefits, security and privacy concerns arise as user's sensitive data are susceptible to both inside and outside attacks.In the proposed system if a file has been uploaded by a user with a duplicate hash value then file will be uploaded on cloud only if he have file and privilege.

Keywords*— *Deduplication, Access control, Hybrid cloud

I. INTRODUCTION

The world moves to digital storage for repository functions, there is increasing demand for systems which will offer secure information storage in an exceedingly cost-efficient manner. Deduplication can yield price savings by increasing the utility of a given quantity of storage. The critical challenge with cloud storage is how to reduce the overload of data storage .so we have come up with deduplication technique .

Deduplication is the process of eliminating the duplicate copy,a pointer is created to reference a data that is backed up. In [8] Deduplication technology typically divides data sets into smaller chunks and uses algorithms to assign each data chunk a hash identifier, which it compares to previously stored identifiers to determine if the data chunk has already been stored.deduplication can take place at file level or sub file level .at file level it detects redundant data within and across files.at sub level it removes redundant copies of identical files.Deduplication technology offers storage and backup administrators a number of benefits, including lower storage space requirements, more efficient disk space used. We have developed a hybrid cloud where we have a private cloud and public cloud. the private cloud is concerned as a proxy to permit information owner/users to firmly perform duplicate talk to differential privileges and also the information is encrypted.the encrypted information is stored in public cloud.

In[4] order to have a secure storage of deduplicated data over a cloud computing. we use the encryption/decryption technique. Encryption is the process of converting a plaintext into a ciphertext. Decryption technique is the process of converting a ciphertext into plain text.the process of symmetric key encryption where same key is used for both encryption and decryption .the key ought to be sending firmly over a network. To stop unauthorized access, a secure proof of possession protocol is additionally required. To produce the proof that the user so owns the similar file when a duplicate is found. Once the proof, sequent users with the similar file area unit reaching to be provided a pointer from the server whereas not having to transfer the similar file.

II. PROBLEM STATEMENT

Aiming at efficiently solving the problem of deduplication with differential privileges in cloud computing, we consider a hybrid cloud architecture consisting of a public cloud and a private cloud. Unlike existing data deduplication systems, the private cloud is involved as a proxy to allow data owner/users to securely per-form duplicate check with differential privileges. Such architecture is practical and has attracted much attention from researchers. The data owners only outsource their data storage by utilizing public cloud while the data operation is managed in private cloud.

III. PROPOSED ARCHITECTURE

The main aim of this architecture is to reduce the overload in cloud and to increase the speed of access. Where we use the deduplication technique, although data deduplication brings a lot of benefits, security and privacy concerns arise as user's sensitive data are susceptible to both inside and outside attacks .so we provide a security to data by access control technique this logic is applied in private cloud. We can have a backup of data in secondary storage but only important data.backup is like having a Xerox of document .we get Xerox of important document ,where as we don't get the Xerox of movie ticket which is not necessary in order to avoid such unwanted data we have comeup deduplication technique.

Aiming at efficiently solving the problem of deduplication with differential privileges in cloud computing. Unlike existing data deduplication systems, the private cloud is involved as a proxy.

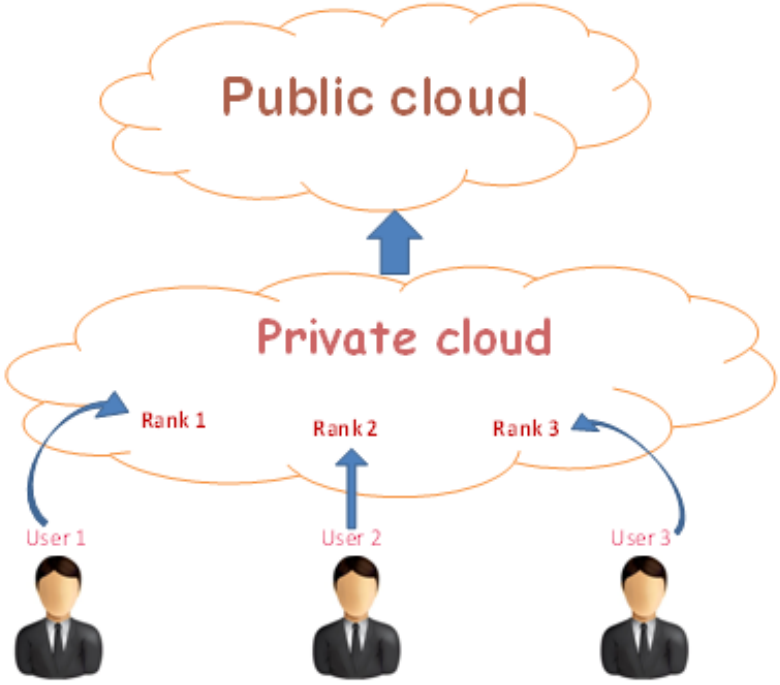

Figure-1 Storage of secure deduplicated data using access control concept.

In the above architecture Figure-1 it consist of public cloud , private cloud and set of users.user1 is given a rank1,based on the rank the authorization permission is given .rank1 user has all the authorization read,write,and duplicate.rank2 user has less privilege then rank1,so rank2 user has read and write permission.rank3 user has very less privilege so he has only read permission.private cloud where user check for the duplicate check before actually storing data in the public cloud.then a data is encrypted in the private cloud and then a rank is provided to user based on the user authorization.later the encrypted data is stored in the public cloud.

Example lets consider we want t store a college database in cloud. We have three user.first user is principal, second user is department head (HOD) and third user is teachers. Principal is given rank1 ,where all the authorization is given he/she can read ,write and duplicate the data .HOD is given rank2 ,less privilege is given he/she can read and write permission is given.teachers is given rank3,where he/she can only read the data cannot make any changes to student database.

Cloud storage are very useful for the corporate as well as for individuals. There may be chance of placing unnecessary duplicate files in cloud storage. Some time, duplicate file placement is needed to overcome the faults in the cloud storage. To overcome this contradiction this Authorized Deduplication system is proposed with security measures.

IV. RESULT

Home page

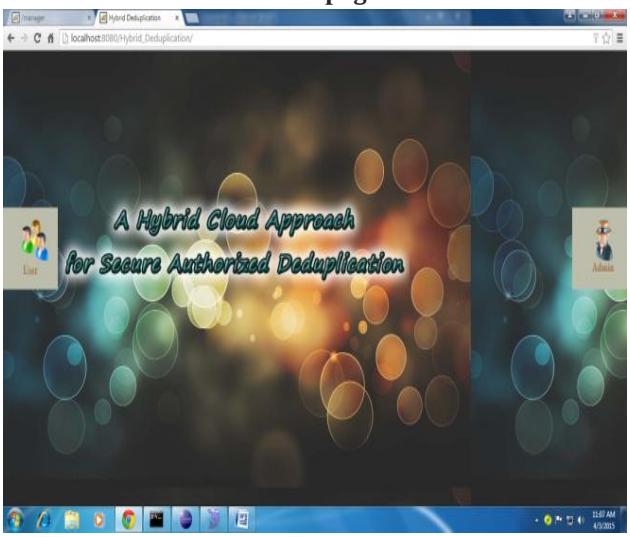

Fig 2.home page

In the above figure 2. we have two models admin and user models.Admin maintence the organization, users are added by admin.users uploads the data and downloads the data.

Admin Home page

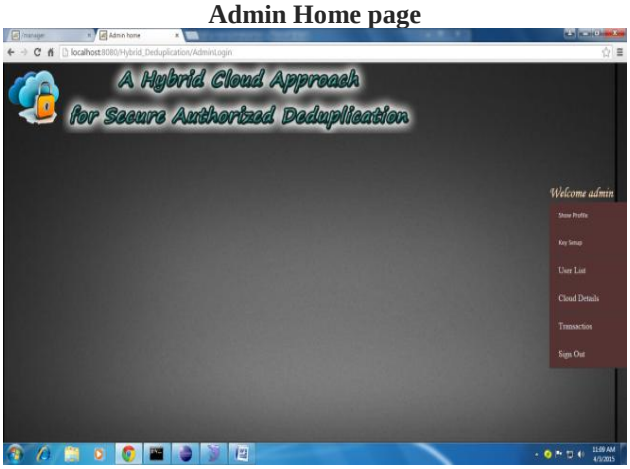

Fig 3.Admin Home page

The above figure 3 is the admin page where we can view the admin profile, we can set up the keys, we can see the user list, we can view the cloud details, even we can view the transactions of all the user details.

Admin Profile

Figure 4- Admin Profile

The above figure 4 is a admin profile .when we click on the show profile the profile details are seen where we can see admin ID No, name, mail id, cell no and address .

Add new admin user

Figure 5 Add new admin users

The above figure 5 is user registration page, where admin has that authority to add the user and rank is also provided by the admin.once all the details are filled the userId the sent to the user mail where user can login with that userId.user can login with Id he/she can upload or download the data

Key setup

Figure 6- Key setup

The above figure 6, where we setup the keys .the keys are the number of users that admin wants to add.

Cloud Details

Figure 7- Cloud Details

The above figure 7,where we can view the cloud details that is Private cloud details and Public cloud details and the cloud status that is active or not.

Transaction Details

Figure 8- Transaction Details

The above figure 8, shows the user transaction details when they have logged in date, time and which file is uploaded at what time, date.

Member Profile Page

Figure 9- Member Profile Page.

The above figure 9, is member profile. When the admin the user their add the user Id is sent to their mail With that Id they can login and they can view the Profile, upload the data, and download the data.

User Upload File

Figure 10- User Upload File

The above figure 10, shows file select from the system to cloud. Click on upload file then choose the file which we want to upload .click submit the file gets uploaded .

User Download File

Figure 11- User Download File

The above figure 11, shows that user can download the file by clicking on select file and then click download, the file gets downloaded.

V. CONCLUSION AND FUTURE WORK

In this paper we have explained how efficiently we can reduces the cloud storage overload by deduplication technique .we have proposed an architecture using hybrid cloud concept.where we have provided the security for the deduplicated data by access control technique, only the authorized person has the permission to deduplicate the data. future work can be when we store the duplicated data as a backup it is limited for the only one or two times.we can duplicate the important document any number of times by load balancing technique so that it is fast to access by any number of clients at a time.

ACKNOWLEDGEMENT

I express my gratitude to **Ms Shwetha K S,** Assistant Professor, Dept. Of ISE, New Horizon College of Engineering, Bangalore for constantly guiding the development of the project. I also thank **Mr. Lokesh M R**, Assistant Professor, Dept Of ISE, New Horizon College of Engineering, Bangalore ,for his valuable suggestions which were the motivating factors in completing the work.

REFERENCE

[1] A Hybrid Method to Reduce Congestion and Provide Data Security in Cloud . Computing.Md Asif Mushtaque, Harsh Dhiman,Shahnawaz Hussain,Pinki .

[2] Data Partitioning Technique to Improve Cloud .Data Storage Security. Swapnil V.Khedkar , A.D.Gawande .

[3] Data Partitioning Technique to Improve Cloud Data Storage Security. Swapnil V.Khedkar , A.D.Gawande .

[4] Reliability and Security of Large Scale Data Storage in Cloud Computing. *C.W. Hsu C.W. Wang Shiuhpyng Shieh* .

[5] An Efficient And Secure Protocol For Ensuring Data Storage Security in Cloud Computing .Syam Kumar P, Subramanian R.

[6] Security Aspects Of Multi-Cloud Computing In Modern Computing Environment Swapnila Mirajkar, Santoshkumar Biradar .

[7] Performance Comparison Of Deduplication Techniques For Storage In Cloud Computing Environment Deepu S R,Bhaskar R,Shylaja B S .

[8] Application-Aware Client-Side Data Reduction and Encryption of Personal Data in Cloud Backup Services.Yin-Jin Fu1, Nong Xiao, Xiang-Ke Liao, *Member*, Fang Liu.

[9] A Policy-based De-duplication Mechanism for Encrypted Cloud Storage. Chuanyi LIU 1, Yancheng WANG2, Jie LIN.

[10] Message-Locked Encryption for Lock-Dependent Messages.Mart.n Abadi, Dan Boneh, Ilya Mironov,Ananth Raghunathan, and Gil Segev2.

[11] RevDedup: A Reverse Deduplication Storage System Optimized for Reads to Latest Backups. Chun-Ho Ng and Patrick P. C. Lee.

[12] Server-Aided Encryption for Deduplicated Storage. Mihir Bellare *Universit*, Sriram Keelveedhi, *San Diego*, Thomas Ristenpart.

[13] Secure Data Deduplication.Mark W. Storer Kevin Greenan Darrell D. E. Long Ethan L. Miller.

[14] International Journal of Computer Engineering and Technology (IJCET), ISSN 0976- 6367(Print), ISSN 0976 – 6375(Online) Volume 4, Issue 4, July-August (2013), © IAEME .

[15] Distributed Key Generation for Secure Encrypted Deduplication.Yitao Duan, NetEase Youdao. Chinese University of Hong Kong May 23, 2014 .

