ISSN: 2321-7782 (Online)

Volume 3, Issue 3, March 2015

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Survey Paper / Case Study

Available online at: www.ijarcsms.com

Authorized Intelligent Data Compression Technique using Hybrid Cloud Approach in Secured way

Pooja S Dodamani[1]
M. Tech Student,
Dept. of Computer science and Engineering,
NMAMIT, Karnataka, India

Pradeep Nazareth[2]
Assistant Professor,
Dept. of Computer science and Engineering,
NMAMIT, Karnataka, India

Abstract: Cloud computing provides lots of services to users and storage benefits to users. Since there is demand for storage in cloud we need to manage the ever-increasing data. So we proposed an Authorized intelligent data compression technique known as deduplication which eliminates the duplicate copies of data in cloud and store only unique files. Since there is lot of issues with quality of service and technical failure occurring the cloud can migrate the data from one cloud to another by providing security.

Keywords: Public cloud, Private cloud, Deduplication, Migration, Users.

I. INTRODUCTION

Cloud computing provides us unlimited resources as services to users over the internet by hiding the implementation and platform details. Cloud service providers provide both parallel computing resources and available storage at low cost. As increasing, amount of data tend to be stored in cloud and provided users who have specified privileges. As the stored volume of data is increasing in the cloud, we need to manage the stored data. To manage the data in cloud computing, the technique called Deduplication is being used which has recently gained more attention.[1]

Data deduplication is a well know intelligent data compression technique used for eliminating the duplicate copies of redundant data in the storage. It is the best technique which can be used for improving the storage utilization and also used to reduce the number of bytes transferred over the network. Instead of keeping the redundant copies of same data it eliminates and keeps only one copy of the physical data and referring the other duplicate data to that copy. Since intelligent data compression technique has lot of benefits, it is sensitive to both outsider and insider attacks. In the traditional encryption system the deduplication process is very difficult in aspects of confidentiality because the users had to encrypt data with their own keys that will provide the different cipher texts for the same identical two copies of data that will make deduplication impossible to be carried out.[2]

Convergent encryption technique has been proposed that will provide confidentiality which makes deduplication feasible. It uses convergent key to encrypt/decrypt the data copy which is obtained after computing the value cryptographic hash of the content. After the key is generated, the users will retain keys and later send only the cipher text of the content to the cloud. The same copy of the data will produce same convergent key that will make deduplication possible. Since to prevent attacks and unauthorized access, the secure proof of ownership concept is provided to authenticate the users. The file can be decrypted only with the convergent key corresponding to Data Owner. Thus the proof of ownership will prevent unauthorized users to access and convergent encryption will allow the deduplication to perform on cipher texts. Each of the file uploaded to the cloud has set of privileges that will allow specifying the access rights and duplicating check to be performed by users. The users who have the particular privileges only will be allowed to carry out the duplication check for the files stored in the cloud. For example- In company there are many different types of privileges assigned to the employees to efficiently manage the stored data i,e stored in cloud by the cloud service providers. The privacy is preserved only by providing services to the authorized users to perform deduplication.

Aiming at solving the problem of deduplication, the hybrid cloud architecture is used where public cloud is used to outsource data and private cloud to manage sensitive operations. The users with specified privileges are only allowed to perform deduplication. By implementing proposed scheme provides authorized duplicate check which provides minimal overhead on upload and deduplication operations.[1]

The cloud computing success is due to customer's ability to use the services on demand with pay-as-you go model that is more convenient. High flexibility and Low costs make migrating to the cloud more compelling. Despite of its advantages, companies tend to hesitate to "move to the cloud," because of concerns which are related to availability of service, legal uncertainties ,data lock-in. Lock-in is particularly problematic even though availability of public cloud is generally high, problems still occur. Company Businesses locked into such type cloud are essentially at a standstill unless and until the cloud is back online. Public cloud providers don't guarantee some required service level agreements i,e businesses locked into such cloud have no guarantees of continuing to provide the required quality of service . Finally, most of the public cloud providers tend to unilaterally change the pricing at any time of the services. So the business locked into such cloud has no mid or long term control over their own IT costs. To address the problems we identified a need for businesses to permanently keep monitoring the cloud which they are using and then able to rapidly migrate to a different cloud, if they find problems or if they estimates predict future issues. To further not complicate the situation, most of the companies not only build on public clouds for all their cloud computing needs but also combine public offerings of cloud with their own private clouds by leading to hybrid cloud setups. Thus we introduced the concept of Meta cloud that will incorporate design time and runtime components. Thus Meta cloud would move away from the existing offerings of technical incompatibilities, thus mitigating the vendor lock-in and It helps the users find right set of cloud services for their particular use case and later supports the application initial deployment and runtime migration.[3]

II. Related Work

Data deduplication is major data compression technique which helps in eliminating the duplicate copies of data that saves the bandwidth and reduces the amount of storage space. Convergent encryption technique is used to provide confidentiality of data and carry out deduplication with authorized duplicate check that incurs minimal overhead, Jin Li, Yan Kit Li, Xiaofeng Chen, Patrick P. C. Lee, Wenjing Lou .[1]

As more private and corporate users outsource data to the cloud service providers, the end-to-end encryption is prominent requirement. A novel idea has been proposed which differentiates the data based on the popularity. An encryption scheme which provides security for unpopular data and weaker security, bandwidth benefits and better storage for popular data. Here deduplication can be done on popular data and secure encryption which protects unpopular data. Here secure scheme is under the Die-Hellman, J. Stanek, A. Sorniotti, E. Androulaki, and L. Kencl[4]

Cloud storage is the service model that makes enterprise and individuals to outsource their data to remote cloud providers with low cost. The service providers must ensure security for the data stored. So in this paper Fade version is used which follows standard version-controlled design for backup that will eliminate the duplicate copies of data among different versions of the backup. Cryptographic protection is applied in order to protect the data backups. It deletes the copy of the data in the same version and makes it not accessible to users and keeps the common data in the version changed with no data modified. This method occurs minimal overhead compared to traditional cloud backup services which do not support assured deletion, A. Rahumed, H. C. H. Chen, Y. Tang, P. P. C. Lee, and J. C. S. Lui[5].

The explanation of attacks on large number of identity based identification security proofs and signature schemes that are implicitly and explicitly in existing literature. So keeping these frameworks how these schemes are derived and enable modular security analysis to understand and unify and simplify the previous work. M. Bellare, C. Namprempre, and G. Neven [6].

Architecture is proposed to deal with the trusted and untrusted commodity when the data is outsourced on the cloud. In the approach users performs critical and sensitive operations on the trusted cloud that will encrypt and verify the data which is stored and less-critical setup phase operation with untrusted cloud, S. Bugiel, S. Nurnberger, A. Sadeghi, and T. Schneider[7].

Lock-in is problematic even when availability of public cloud is generally high, problems still occur. Public cloud providers don't guarantee some required service level agreements i,e businesses locked into such cloud have no guarantees of continuing to provide the required quality of service .To address the problems we identified a need for businesses to permanently keep monitoring the cloud which they are using and than able to rapidly migrate to a different cloud, if they find problems or if they estimates predict future issues Benjamin Satzger, Waldemar Hummer, Christian Inzinger, Philipp Leitner, and Schahram Dustdar[3].

III. System Model

This paper proposes the IND-OCPA-P model to analyze the security of the proposed EOB and the encryption schemes supporting an efficient range query over encrypted data.

Architecture

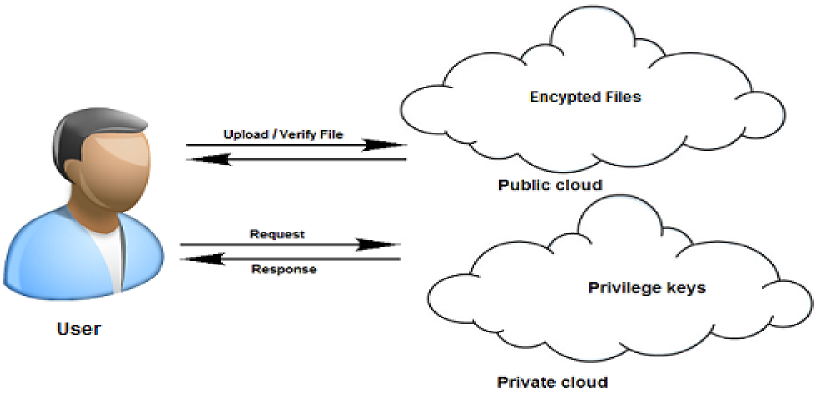


Figure 1. Architecture diagram

The architecture is for enterprise network that consists of a group of clients who will use the deduplication technique which storing the data that reduces the amount of storage of redundant data. We have three entities in our system i,e users, public cloud, and private cloud.

The access rights are defined only with the set of privileges users who have the necessary privileges will be carrying out the deduplication. The public cloud is where we outsource the data by service provider and that can be publicly accessible and private cloud where users can carry out critical operations of duplicate check for the file.

» Public cloud- This is entity which provides the storage of data service in public cloud. It outsources the data on behalf of users. In order to reduce the storage cost we need to eliminate the redundant data by deduplication and keep only one copy of unique data.

» Data users- A user entity will want to outsource their data on to public cloud and get the access of data later. In the public cloud user only upload unique copy data in order to save bandwidth and storage space. The users who have set of privileges are only allowed to carry deduplication. Since the file will be protected with privileges and convergent key.

» Private cloud- The private cloud is an entity which provides the environment between users and public cloud. The privileges and private keys will be managed by private cloud.

IV. Implementation

In our proposed authorized secure deduplication systems, we model three entities as the separate Java Programs. The client programs that will be used to carry out the File upload process and generation of the signature. Cloud server that manages to store the data on the cloud server with the encrypted keys and signature. End users who have specified privileges can be able to decrypt the data and get the file. The end user who has specified privileges to access the files.

Our implementation provides the following functions

» ***Client program***

- The users with specified set of privileges will be able to browse the required file.
- The signature is being generated by computing SHA-1 hash of the file.
- The file is encrypted using AES algorithm and uploaded to the cloud server.
- The users with privileges can also carry out the deduplication check on files stored on the cloud server.

» ***Cloud server***

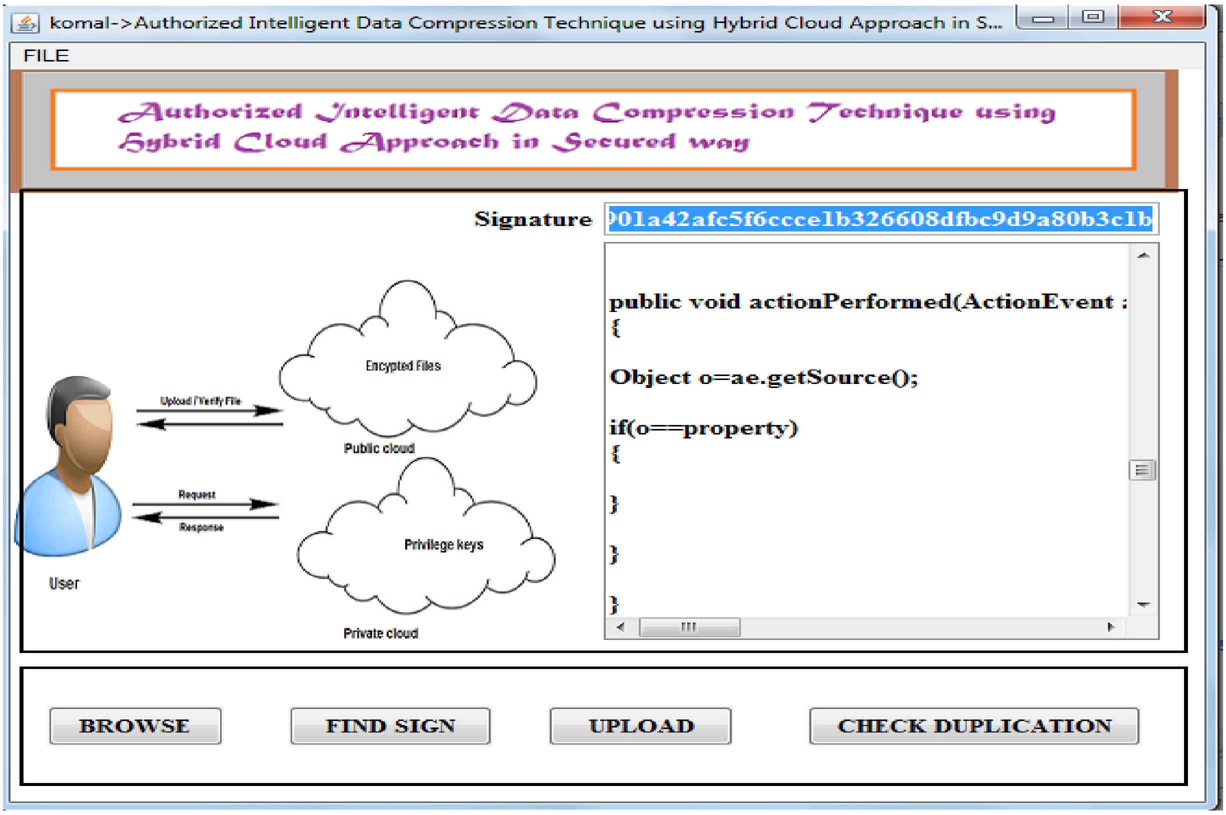
- Cloud server is the service provider where the files of the users are uploaded. The users who are having privileges given by cloud server are only able to upload and carry out deduplication check. The users trying to upload duplicated files will be blocked.
- The users can carry out duplication check each time before uploading the file on the cloud servers. Same set of files generate same convergent key making deduplication easy.
- The users willing to view the files need the privileges to be provided by cloud server and then access the required files.
- Migration of data from one cloud server to different cloud server in case of technical issues occurred.

» ***End user***

- The users who have the privileges of the cloud server will be able to download with the specified secret key and filename generated during encryption and by having it, the file will be decrypted and provided to the users.

V. Evaluation

1. Browse the file to be uploaded and check deduplication.
2. Generate signatures for the file.
3. Upload the required file to cloud servers.
4. Check for duplication of files after the signature generation by comparing.



komal->Authorized Intelligent Data Compression Technique using Hybrid Cloud Approach in S...

Authorized Intelligent Data Compression Technique using Hybrid Cloud Approach in Secured way

Signature 01a42afc5f6ccce1b326608dfbc9d9a80b3c1b

```
public void actionPerformed(ActionEvent :
{

Object o=ae.getSource();

if(o==property)
{

}

}

}
```

Figure 2 . Generation of signature

VI. Conclusion

In this paper, the secure authorized intelligent data compression technique was proposed to protect the security of data by having differential privileges given to users during duplication check. The generation of the same convergent key for two identical copies of data helps us to easily carry out the deduplication. Migration of data from one cloud to another will help us to secure the data during data overhead occurred in cloud server. We implement our proposed secure authorized intelligent data compression scheme and conducted testbed experiments. We showed that our scheme incurs minimal overhead.

References

1. Jin Li, Yan Kit Li, Xiaofeng Chen, Patrick P. C. Lee, Wenjing Lou. A Hybrid Cloud Approach for Secure Authorized Deduplication. IEEE Transactions on Parallel and Distributed Systems,Volume:PP,Issue:99,Date of Publication :18.April.2014
2. Danny Harnik, Benny Pinkas, Alexandra Shulman- Peleg "Side Channels in Cloud Services Deduplication in Cloud Storage.‖ COPUBLISHED BY THE IEEE COMPUTER AND RELIABILITY SOCIETIES, NOVEMBER/DECEMBER 2010.
3. Benjamin Satzger, Waldemar Hummer, Christian Inzinger, Philipp Leitner, and Schahram Dustdar. Winds of Change: From Vendor Lock-In to the Meta Cloud. Published by the IEEE Computer Society,2013
4. J. Stanek, A. Sorniotti, E. Androulaki, and L. Kencl. A secure data deduplication scheme for cloud storage. In Technical Report, 2013.
5. 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.
6. M. Bellare, C. Namprempre, and G. Neven. Security proofs for identity-based identification and signature schemes. J. Cryptology, 22(1):1–61, 2009.
7. 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.

AUTHOR(S) PROFILE

Pooja S Dodamani, received Bachelor degree B.tech from Vishweshwaraya Technological University,Belgaum. She is now pursuing Masters Degree Mtech computer science and engineering department at NMAM Institute of Technology, Nitte.

Pradeep Nazareth, received Bachelor degree B.tech from Vishweshwaraya Technological University,Belgaum, and Masters Degree Mtech from SJCE Mysore. He is now working as Assistant Professor in computer science and engineering department at NMAM Institute of Technology, Nitte.