Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 5.258

IJCSMC, Vol. 5, Issue. 4, April 2016, pg.229 – 235

An Efficient and Secure Deduplication Management in Private Cloud

[1]Revathi.K, [2]Golda Selia.I

[1]M.Tech Student, Department of Computer Science & Engineering, Dr. M.G.R. Educational and Research Institute, DR MGR University, Chennai, Tamil Nadu, India

[2]Assistant Professor, Department of Computer Science & Engineering, Dr. M.G.R. Educational and Research Institute, DR MGR University , Chennai, Tamil Nadu, India

[1] Revathimtech2014@gmail.com, [2] Goldaselia@gmail.com

Abstract- The storage space and the bandwidth in the cloud storage can be decreased by using the techniques for eliminating the data duplication. The repeating data or the duplicate data occupies a large space in the cloud and also its bandwidth. Before the transmission of the sensitive data the encryption techniques are to be used to protect the integrity and the confidentiality of the data supporting the deduplication at the same time by using the Filetoken for the cloud data. The file tokens are used for encrypting the data and avoiding the redundancy of the data. The problems that are raised due to deduplication and protect the data security this paper makes the effort to address them. Apart from the traditional checks for the deduplication the users with differential privileges are taken in the duplicate along with the data itself. In this paper we propose the hybrid cloud architecture supporting the checking of the duplication in a cloud storage. The security model describes that our process of security analysis and the security definitions are clearly explained. Here we introduce a proof of identity that we are going to implement for our authorized duplicate check scheme. We present that our process is authorized check for the duplicate scheme that incurs minimal when compared to the normal operations.

Index Terms—Filetoken, Deduplication

1. INTRODUCTION

The virtualized sources that are in the cloud computing provides us a great service across the internet, while making the platform as well as the implementation details of them too a secure aspect. Today cloud have been raised to a good volume and provides us a good storage and the computing parallel at very low cost relatively. There is an increasing amount of cloud storage due to its security as well as its volume. But there have been a greater discussion on the privileges on the user as well as the deduplication of the data across the cloud.

To make a scalable and a trustable cloud computing, the deduplication has been a well-known technique turning its attention gradually. The deduplication is nothing but the specialized way of data compression for the elimination of the data deduplication. This typical technique is used for the transmitting the data over network securely and maintaining the data with a single copy of the data, rather than multiple copies of data being used. This reduces the redundant data to a huge extent. While uploading or downloading the token request is to be made and then the file token is generated by a private cloud. Then he user can be able to upload or download the data for the user in a secure manner. This enhances the duplication as well by converting the data into bytes and then comparing the data for the ease of making the physical data distinct.

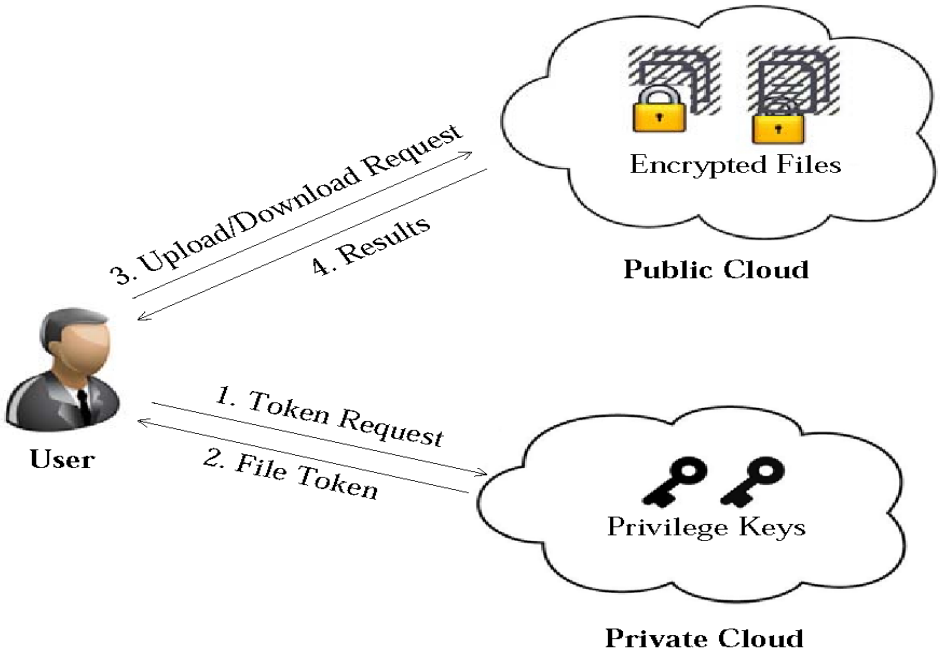

Fig 1: Data transferring with file tokens

a) Authorization control creation and Key Generation:

The private keys for privileges will not be issued to users directly, which will be kept and managed by the private cloud server instead. In this way, the users cannot share these private keys of privileges in this proposed construction. The privilege key sharing among users in the above straightforward construction.

To get a file token, the user needs to send a request to the private cloud server. The process of managing the cryptographic keys in the cryptosystem includes in dealing with the many scenarios ,example of them include the way of using, storing and the key replacement. There are many protocols that are used to include the cryptographic way of sending and receiving the data some of them include the designing of the protocol and the procedure the user uses and the key generated in the servers.

The cipher key is concerned for the user at every level of transmission and receiving. In the contrast for the scheduling of the key that is typically for handling the material operations for the cipher key.

The key management is successful in all the level of the security for the cryptosystem. The system policy is the difficult aspect for the cryptography because it involves the many factors right from the system policy to the user training and the organization and departmental coordination between all these elements.

b) Owner uploading in Private Cloud:

A user can upload a file and stored in a private cloud. The private keys for privileges will not be issued to users directly, which will be kept and managed by the private cloud server instead. In users cannot share these private keys of privileges in this proposed construction The privilege key sharing among users in the above straightforward construction. The authorized duplicate check for this file can be performed by the user with the public cloud before uploading this file.

The cloud storage proposed in our system is the combination of different clouds like the private and public also known as the community cloud that enables the users to use them. But in reality they are bound together for providing a good service of the data storage. These are dedicated for the cloud resources to connect the collocation with a managed and dedicated servers.

The private cloud service can be defined as cloud computing service that they are integrated with the different types of the cloud servers that are available like the public ,private and the community cloud servers. The separate hybrid cloud service that is used in the boundaries of the given category of browsers enables us to extend their volumes in terms of the cloud services, the integration and the customization with the cloud service providers.

There is a varied use of the case of the hybrid cloud composition. If we consider the organization that transfers the sensitive data in a local home network through the private cloud application can be made by using the business intelligence provided on the public cloud as well as a software service. These transmission depends on many number of factors for organizing the quality data transfer in a good fashion.

There is a wide range of cloud users increasing graph. This scenario makes the data to be transferred at a high rate and at a high speed. These capabilities enables the cloud to employ the data to a good turning in terms of the private or public cloud. The most important features that are used for this functionality is the bursting tetchiness where the data is so secure and the breached in the flow are never allowed. The prior advantage of the hybrid cloud is it has the elements that allow the users to make them more compatible for the user requirement. This creates a good amount of demand in the IT sector as well as the raising of the converting the complex task to a simple and to a secure way.

c) Detect Deduplication:

Convergent encryption provides data confidentiality in deduplication. A user derives a convergent key from each original data copy and encrypts the data copy with the convergent key. In addition, the user also derives a tag for the data copy, such that the tag will be used to detect duplicates. To detect duplicates, the user first sends the tag to the server side to check if the identical copy has been already stored.

Data deduplication is an enhanced technique for the data compression as well as the duplication of the data for repeating the copies to prevent it from the unsecured way of using them. This improves the utilization of the data transfers used in the cloud. There are single instance data storage measures that are taken from the initial stage of data being processed. In the

deduplication process the data is being processed by using chunks of data and being transferred for the analysis and then they are given with the redundant chunk that remove them.

In this way the copy of the stored data when ever matched then the chucks are reflected back explaining that the data is being provided and cannot be replaced or any other factors for the improvement of the data process.

d) Key Exchanging:

The private keys for the privileges are managed by the private cloud the file token requests from the users. The interface offered by the private cloud allows user to submit files. In queries to be securely stored and computed respectively. The private cloud server will also check the user's identity before issuing the corresponding file token to the user. The authorized duplicate check for this file can be performed by the user with the public cloud before uploading this file. The data is being spoofed first and then the data is being compared for the storing. Before this process the data looking for the duplication is no need any wait for the calculation or the lookups that are being made or completed for a great extent of the data transferred. There by this ensures that the data is not being transferred to a great extent. The active files are compared and then they are processed for the transmission of the data.

e) Verification and File Retrieving:

The private cloud server will also check the user's identity before issuing the corresponding file token to the user. An identification protocol use the proof and verification algorithm respectively. It also initializes a PoW protocol POW for the file ownership proof. In file retrieving it first sends a request and the file name to the S-CSP. Upon receiving the request and file name, the S-CSP will check whether the user is eligible to download file.

The data that is been deduplicated can be read back if one of the file is delivered in chunks. This can simply produce the link with the referred chunks of data. All these are ended with the transparent transfer to the end application user.

For example the chunking between the various factors like the technology, implementation and the data factors. There are some physical layer constrains that are removed by the deduplication techniques used by us. The general method for the intensive chunking is the sliding-block.

Architecture Diagram:

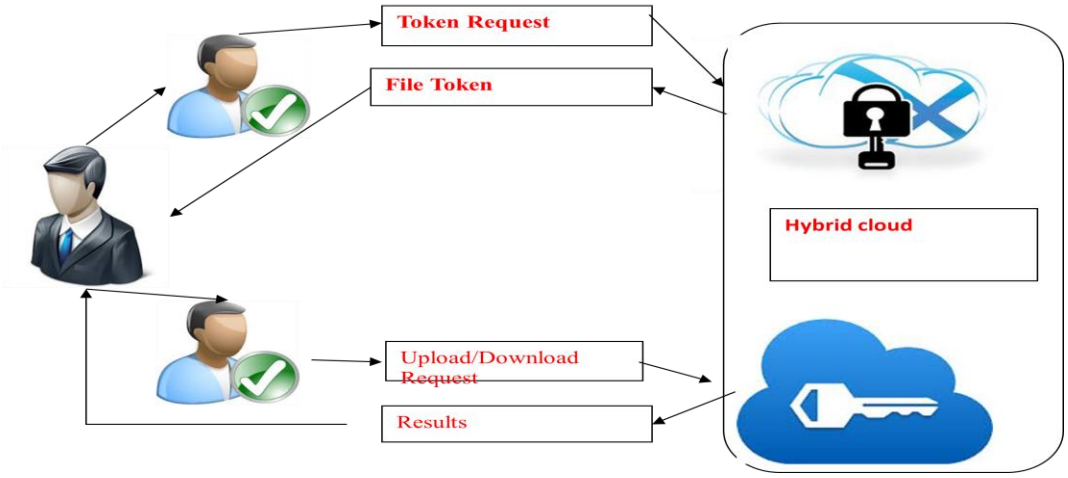

Fig:2 - Data Depulication by using file keys generation

2. CONCLUSION

In this paper, we have tested the data with different access privileges and data is processed with different techniques like the data dedupication check supporting the data compression and saving bandwidth and the volume of the cloud storage. We also presented several new deduplication constructions supporting authorized duplicate check in private cloud architecture, in which the duplicate-check tokens of files are generated by the private cloud server with private keys. Security analysis demonstrates that our schemes are secure in terms of insider and outsider attacks specified in the proposed security model. As a proof of concept, we implemented a prototype of our proposed authorized duplicate check scheme and conduct test based experiments on our prototype. We showed that our authorized duplicate check scheme incurs minimal overhead compared to convergent encryption and network transfer.

REFERENCES:

1. M. Bellare, S. Keelveedhi, and T. Ristenpart. Dupless: Serveraided encryption for deduplicated storage. In USENIX Security Symposium, 2013.
2. M. Bellare, S. Keelveedhi, and T. Ristenpart. Message-locked encryption and secure deduplication. In EUROCRYPT, pages 296–312, 2013.
3. OpenSSL Project. http://www.openssl.org/.
4. P. Anderson and L. Zhang. Fast and secure laptop backups with encrypted de- duplication. In *Proc. of USENIX LISA*, 2010.
5. M. Bellare, S. Keelveedhi, and T. Ristenpart. Dupless: Server- aided encryption for deduplicated storage. In *USENIX Security Symposium*, 2013.
6. M. Bellare, S. Keelveedhi, and T. Ristenpart. Message-locked encryption and secure deduplication. In *EUROCRYPT*, pages 296312, 2013.
7. M. Bellare, C. Namprempre, and G. Neven. Security proofs for identity-based identification and signature schemes. *J. Cryptology*, 22(1):1-61, 2009.
8. 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.
9. 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.
10. 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.
11. D. Ferraiolo and R. Kuhn. Role-based access controls. In *15th NIST-NCSC National Computer Security Conf.*, 1992. GNU Libmicrohttpd.

