International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

Secure Authorized Deduplication in Hybrid Cloud

Dr. K.V. Reddy[1]; T.S.Srinivas [2]& J.Gangadhar[3]

[1]Professor, Dept of CSE, Marri Laxman Reddy Institute of Technology & Management, Hyderabad, Telangana

[2] Associate Professor, Dept of CSE, Marri Laxman Reddy Institute of Technology & Management, Hyderabad, Telangana

[3]M.Tech CSE (PG Scholar), Dept of CSE, Marri Laxman Reddy Institute of Technology & Management, Hyderabad, Telangana

ABSTRACT

Cloud computing has showed up as a popular design in managing world to back up managing large volumetric details using cluster of commodity computer systems. It is the newest effort in offering and managing computing as a service. Either program or Application, it is used to describe both. A cloud computing paradigm dynamically assigns, configures, relocates and de provisions these computing resources as needed. It also describes applications that are to be extended accessible through the Internet. Data deduplication is a technique for reducing the amount of storage space an organization needs to save its data.

In most organizations, the storage systems contain duplicate copies of many pieces of data. For example, the same file may be saved in several different places by different users, or two or more files that aren't identical may still include much of the same data. Deduplication eliminates these extra copies by saving just one copy of the data and replacing the other copies with pointers that lead back to the original copy. Companies frequently use deduplication in backup and disaster recovery applications, but it can be used to free up space in primary storage as well. To avoid this duplication of data and to maintain the confidentiality in the cloud we using the concept of Hybrid cloud. To protect the confidentiality of sensitive data while supporting deduplication, the convergent encryption technique has been proposed to encrypt the data before outsourcing. To better protect data security, this paper makes the first attempt to formally address the problem of authorized data deduplication.

I. INTRODUCTION

Cloud computing has become a necessity today when the company plans to increase capacity "or capabilities on the fly without getting to invest new infrastructure, training new individual purchase new license application, etc. based service encompasses any subscription or pay per use which extends the existing IT capabilities of the company, current time through Online. To make data management scalable in cloud computing, deduplication has been a well-known technique and has attracted more and more attention recently. Data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data in storage. The technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. Instead of keeping multiple data copies with the same content, deduplication eliminates redundant data by keeping only one physical copy and referring other redundant data to that copy [1].

In order to save cost an efficiently management , the data will be moved to the storage server provider in the public cloud with identified privileges the deduplication techniques will be applied to store only in copy of the same file. Because of privacy concern, some files will

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

be encrypted an acceptable the duplicate check by employee with specified privileges to realize the access control [2]. Traditional, deduplication system based on merging cipher, although given that privacy to certain amount; do not provision the replacement check with variance privileges in other word, no differential privileges have been considered in deduplication based on convergent encryption technique. It s seems to be reversed if we want to recognize both deduplication and differential authorization duplicate check at the same time.

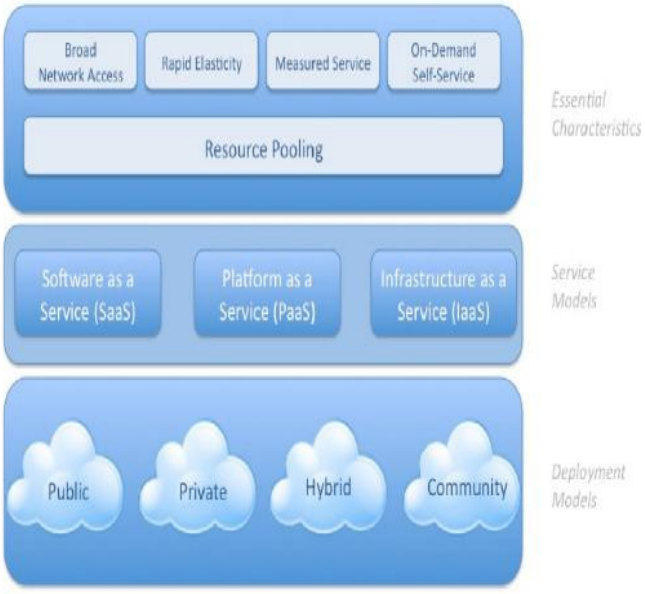

Figure 1: Cloud Computing Architecture

In computing, data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data. Related and somewhat synonymous terms are intelligent (data) compression and single-instance (data) storage. This technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. In the deduplication process, unique chunks of data, or byte patterns, are identified and stored during a process of analysis. As the analysis continues, other chunks are compared to the stored copy and whenever a match occurs, the redundant chunk is replaced with a small reference that points to the stored chunk [3]. Given that the same byte pattern may occur dozens, hundreds, or even thousands of times (the match frequency is dependent on the chunk size), the amount of data that must be stored or transferred can be greatly reduced.

Keeping data backups permanently is undesirable, as sensitive information may be exposed in the future because of data breach or erroneous management of cloud operators. Thus, to avoid liabilities, enterprises and government agencies usually keep their backups for a finite number of years and request to delete (or destroy) the backups afterwards. For example, the US Congress is formulating the Internet Data Retention legislation in asking ISPs to retain data for two years, while in United Kingdom; companies are required to retain wages and salary records for six years.

A Hybrid Cloud is a combined form of private clouds and public clouds in which some critical data resides in the enterprise's private cloud while other data is stored in and accessible from a public cloud. Hybrid clouds seek to deliver the advantages of scalability, reliability, rapid deployment and potential cost savings of public clouds with the security and increased control and management of private clouds. As cloud computing becomes famous, an increasing amount of data is being stored in the cloud and used by users with specified privileges, which define the access rights of the stored data. The hybrid cloud gives the functionality, scalability, reliability, fast deployment and cost saving of public cloud storage by reducing redundancy in data [3, 4].

II.RELATED WORK

Instead of storing data on single server, cloud storage refers to use third party provider. The cloud storage interface is installed based on client requirement to different storage nodes. So

operating in cloud storage is similar to local storage operating device. Using various network devices cloud storage changes the Application Programming Interface (API) of cloud storage (like Simple Object Access Protocol (SOAP), Representational State Transfer (REST)). For achieving good result many of the researchers have suggested to use De-duplication and Feedback control schemes [5].

Ohsaki said that the term job manager, an independent program generated according to different demand and Feedback control system is used from Resource Management Mechanism which is used to distribute resources and to manage Quality of Service. The job manager is generated according to a different demand. Thus as increase in data, multiple job managers might be generated, which increases the server workload [7].

Previous deduplication systems cannot support differential authorization duplicate check, which is important in many applications. In such an authorized deduplication system, each user is issued a set of privileges during system initialization. Each file uploaded to the cloud is also bounded by a set of privileges to specify which kind of users is allowed to perform the duplicate check and access the files. Before submitting his duplicate check request for some file, the user needs to take this file and his own privileges as inputs. The user is able to find a duplicate for this file if and only if there is a copy of this file and a matched privilege stored in cloud. For example, in a company, many different privileges will be assigned to employees [5].

J. Xu proposed growing need for secure cloud storage services and the attractive properties of the convergent cryptography lead us to combine them, thus, defining an innovative solution to the data outsourcing security and efficiency issues. Our solution is based on a cryptographic usage of symmetric encryption used for enciphering the data file and asymmetric encryption for Meta data files, due to the highest sensibility of this information towards several intrusions [6].

In order to save cost and efficiently management, the data will be moved to the storage server provider (S-CSP) in the public cloud with specified privileges and the deduplication technique will be applied to store only one copy of the same file. Because of privacy consideration, some files will be encrypted and allowed the duplicate check by employees with specified privileges to realize the access control. Traditional deduplication systems based on convergent encryption, although providing confidentiality to some extent; do not support the duplicate check with differential privileges [10]. In other words, no differential privileges have been considered in the deduplication based on convergent encryption technique. It seems to be contradicted if we want to realize both deduplication and differential authorization duplicate check at the same time.

1.Fast and secure laptop backups with encrypted de-duplication:

Many people now store large quantities of personal and corporate data on laptops or home computers. These often have poor or intermittent connectivity, and are vulnerable to theft or hardware failure. Conventional backup solutions are not well suited to this environment, and backup regimes are frequently inadequate [12]. This paper describes an algorithm which takes advantage of the data which is common between users to increase the speed of backups, and reduce the storage requirements. This algorithm supports

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

client-end per-user encryption which is necessary for confidential personal data.

2. Message-locked encryption and secure de-duplication:

We formalize a new cryptographic primitive, Message-Locked Encryption (MLE), where the key under which encryption and decryption are performed is itself derived from the message. MLE provides a way to achieve secure deduplication (space-efficient secure outsourced storage), a goal currently targeted by numerous cloud-storage providers [12]. We provide definitions both for privacy and for a form of integrity that we call tag consistency.
Based on this foundation, we make both practical and theoretical contributions. On the practical side, we provide ROM security analyses of a natural family of MLE schemes that includes deployed schemes. On the theoretical side the challenge is standard model solutions, and we make connections with deterministic encryption, hash functions secure on correlated inputs and the sample-then-extract paradigm to deliver schemes under different assumptions and for different classes of message sources. Our work shows that MLE is a primitive of both practical [10].

3. Security proofs for identity-based identification and signature schemes:

This paper provides either security proofs or attacks for a large number of identity-based identification and signature schemes defined either explicitly or implicitly in existing literature. Underlying these is a framework that on the one hand helps explain how these schemes are derived and on the other hand enables modular security analyses, thereby helping to understand, simplify, and unify previous work. We also analyze a generic folklore construction that in particular yields identity-based identification and signature schemes without random oracles [9].

III.Proposed System

In the proposed system we are achieving the data deduplication by providing the proof of data by the data owner.

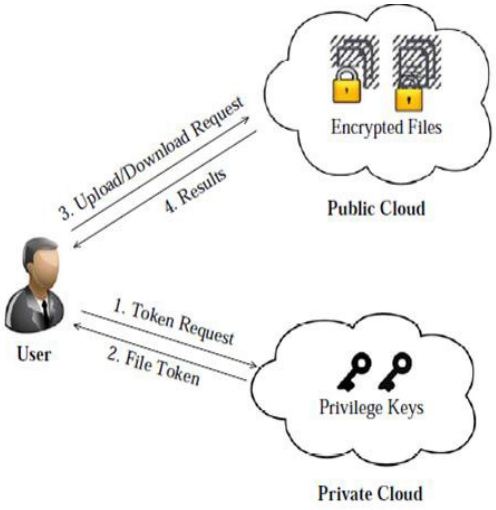

Figure 2: System Architecture

The user needs to take this file and his own privileges as inputs. The user is able to find a duplicate for this file if and only if there is a copy of this file and a matched privilege stored in cloud.

A. Data Encryption of Files:

In this approach we use same secret key k for the encryption as well as decryption of the data. This will help us to convert the plain text to the cipher text for these we have used three basic method. KeyGse: K is the method to create key generation steps that form k using security method 1. EnSE(k,M): C is the symmetric key encryption algorithm that take secret key k and message m and then output the cipher text Dc(K,C): m is symmetric decryption algorithm that use the secret key and cipher text c and the output the original message m.

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 08, August 2015
Available at http://internationaljournalofresearch.org

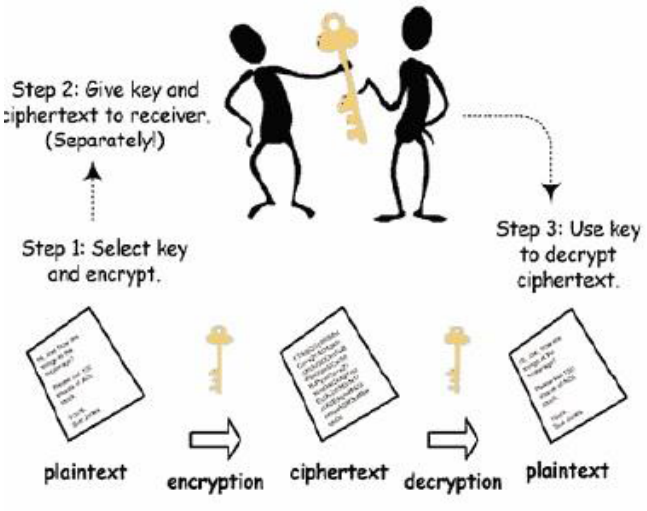

Figure 3: Process of making data confidential

B. Important Confidential Data:

In this method it provide data confidentiality along with encryption method here we use convergent key from each original information with the help of convergent key here user also produced the tag for the information that tag will be used to identify the duplicate copy .

c. Identity of Data:

Here user is supposed to prove that the information which he want to upload or download its self-data.

Advantages of proposed system:

a.The private keys for privileges will not be issued to users directly, which will be kept and managed by the private cloud server.

b.The users cannot share these private keys of privileges in this proposed construction, which means that it can prevent the privilege key sharing among users in the above straightforward construction.

c.To get a file token, the user needs to send a request to the private cloud Server.

d.Extensive security and performance analysis shows that the proposed scheme is highly effective and resilient to malicious data modification attacks.

e. Symmetric encryption keys are provided to encrypt and decrypt the data respectively. Unlike traditional encryption, convergent encryption is more efficient to practice and implement.

IV.Conclusion

cloud computing increasing day by day cloud computing become a research topic for better confidentiality and security aspects we proposed a new de-duplication technique assisting authorized duplicate check in hybrid cloud system in which duplicate check tokens of files are created by private cloud server with private keys .our proposed system include identity of data owner so it will help as handle better security issues in cloud computing. Several new deduplication constructions supporting authorized duplicate check in hybrid 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.

REFERENCES

[1]. Pooja S Dodamani, Pradeep Nazareth, 2014, A Survey on Hybrid Cloud with De-Duplication

[2] Danny Harnik, Benny Pinkas, Alexandra Shulman- Peleg , 2010, Side Channels in Cloud Services Deduplication in Cloud Storage.

[3] Hui Zhang, Guofei Jiang, Kenji Yoshihira, Haifeng Chen and AkhileshSaxena ,2009, A Hybrid Cloud ComputingModel

[4] Borja Sotomayor, Rubén S. Montero and Ignacio M. Llorente, Ian Foster ,2009, Virtual Infrastructure Management in Private and Hybrid Clouds.

[5] David Geer, 2008, Reducing the Storage Burden via Data Deduplication.computer.org.

[6] D. Ferraiolo and R. Kuhn. Role-based access controls. In 15th NIST-NCSC National Computer Security Conf., 1992.

[7] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman Peleg. Proofs of ownership in remote storage systems.

[8] In Y. Chen, G. Danezis, and V. Shmatikov, editors, ACM Conference on Computer and Communications Security, pages 491–500. ACM, 2011.

[9] J. Li, X. Chen, M. Li, J. Li, P. Lee, and W. Lou. Se-cure deduplication with efficient and reliable convergent key management. In IEEE Transactions on Parallel and Distributed Systems, 2013.

[10] M. Bellare, S. Keelveedhi, and T. Ristenpart. Dupless: Serveraided encryption for deduplicated storage. In USENIX Security Symposium, 2013.

[11] M. Bellare, S. Keelveedhi, and T. Ristenpart. Message-locked encryption and secure deduplication. In EUROCRYPT, pages 296ï¿½312, 2013.

[12] M. Bellare, C. Namprempre, and G. Neven. Security proofs for identity-based identification and signature schemes. J. Cryptology, 22(1):1ï¿½61, 2009.

[13] 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.