Survey based on Secure Authorized Deduplication Hybrid Cloud Approach

Backialakshmi.N
ME Student
Department of Computer Science & Engineering
Adhiyamaan College of Engineering, Hosur, Tamil Nadu, India

Manikandan.M
Assistant Professor
Department of Computer Science & Engineering
Adhiyamaan College of Engineering, Hosur, Tamil Nadu, India

Abstract

Data de-duplication is one of important data compression techniques for eliminating duplicate copies of repeating data, used in cloud storage to reduce the amount of storage space and save bandwidth. To protect the confidentiality of sensitive data while supporting de-duplication, the convergent encryption technique has been proposed to encrypt the data before outsourcing. To protect data security, it makes the first attempt to formally address the problem of authorized data de-duplication.
Keywords: Data de-duplication, Convergent Encryption, Cloud Storage, Data compression, Security

I. INTRODUCTION

Cloud computing provides seemingly unlimited "virtualized" resources to users as services across the whole Internet, while hiding platform and implementation details. Today's cloud service providers offer both highly available storage and massively parallel computing resources at relatively low costs. As cloud computing becomes prevalent, an increasing amount of data is being stored in the cloud and shared by users with specified privileges, which define the access rights of the stored data. One critical challenge of cloud storage services is the management of the ever-increasing volume of data.

To make data management scalable in cloud computing, de duplication has been a well-known technique and has attracted more and more attention recently. Data de duplication 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, de duplication eliminates redundant data by keeping only one physical copy and referring other redundant data to that copy. De duplication can take place at either the file level or the block level. For file level de duplication, it eliminates duplicate copies of the same file. De duplication can also take place at the block level, which eliminates duplicate blocks of data that occur in non-identical files.

II. SURVEY PAPER

A. Secure Data De-duplication:

M.W.Storer, K.Greenan, D.D.E.long, E.L.Miller et al, identifying common chunks of data both within and storing in only once, de-duplication can yield cost savings by increasing the utility of a given amount of storage. De-duplication exploits identical content; while encryption attempts to make all content appear random. The same content encrypted with two different keys results in very different cipher text. Combining the space efficiency of de-duplication with the secrecy aspects of encryption is problematic. Developed a solution that provides both data security and space efficiency in single-server storage and distributed storage systems.

Convergent encryption (CE) to enable encryption while still allowing the de-duplication common on chunks. Convergent encryption users a function of the hash of the plaintext of chunks have encryption key this technique dose leak knowledge that a particular cipher text and plaintext. All data chunking and encryption occurs and the client. Plaintext data is never
Transmitted, strengthening the system against both internal and external adversaries.

Finally the map that associates chunks to a given file is encrypted using a unique key, limiting the effect of key compromise to single file. The key are stored within the system in such that users only need to maintain a single private key regardless of the number of files to which they have access.

B. Security Proof for Identity Based Identification and Signature Schemes:

M.Bellare, C.Namprempre, G.Neven et al provides either security proofs or attacks for a large number of identity based identification and signature schemes defined either explicitly or implicitly in existing. This frame works that one hand it helps

explain how the schemes are derived and enables the modular security analysis. IBI is an authority having a master public key and a master secret key.

This authority can provide a user with a secret key based on its identity. The user, playing the role of a prover, can then identity itself to a verifier in a protocol in which the verifier begins by knowing only the claimed identity of the prover and the master key of the authority. IBS scheme is similar except that the user signs messages, rather than identifying itself and verification of a signature requires knowledge only of the identity of the signer and the master public key.

C. Fast and Secure Laptop Backups with Encrypted De-Duplication:

P.Anderson, L.Zhang et al to describes conventional backup algorithm which takes advantages of the data which is common between users in increase the speed of backups and reduce the storage requirement. This algorithm supports client-end- user encryption which is necessary for confidential personal data. It also supports a unique feature which allows immediate detection of common sub trees, avoiding the need to query the backup system for every life.

To descries a prototype implementation of this algorithm for Apple OS X, and present an analysis of the potential effectiveness, using real data obtained from a set of typical users.

De – duplication of a hashing function can be used to return a unique key for a block of data, based only on the contents of the data; if two people have the same data, the hashing function return the same key the keys user index for storing the data block, any attempt to stores multiple copies of the same block will be detected immediately.

Encrypting data invalidates the de duplication, two identical data blocks, encrypted with different keys. The encryption key for the data block is derived from the contents of the data using a function is similar to the has function. Two identical data blocks yield identical encrypted blocks which can de duplicated in the normal way. Each block has a separate encryption key.

D. Twin Clouds: An Architecture for Secure Cloud Computing:

S.Bugiel, S.Nurnberger, A.Sadeghi, T.Schneider et al cloud computing promises a more cost effective enabling technology to outsource storage and computations. For secure outsourcing of data and arbitrary computation are based on a single tamper - proof hardware. The user communicates with trusted cloud which encrypts and verifies the data stored and performed in the untrusted commodity cloud. Secure computation of arbitrary functions on confidential date can be achieved based on fully homomorphic encryption.

The trusted cloud requires only a constant amount of storage and is used constantly in the setup phase for pre- computing encryptions. The trusted cloud is used mostly in the setup phase to encrypt the outsourced data.

E. Proof of Ownership an Remote Storages System:

S.Halevi, D.Harnik, B.pinkas, A.Shulman-Peleg et al ,the promising technologies keeps their cost down is de-duplication, which stores only a single copy of repeating data. Client – side de-duplication attempt to identify de-duplication opportunities at the client and saved the b and with of uploading copies of existing files to the server. To identify attacks that explore client-side de duplciation, allowing and attacker to gain access to arbitrary- size file of other users based on a very small has signature of the files. And attacker knows the has signature of a file can be storage services that it ownes file; the server let the attacker download the entire file.

To overcome the attacks to introduce the notion of proofs- of-ownerships (pows), lets a client efficiently prove to a server that the client holds a file rather than the some short information. The formalize the concept of proof of ownership, under rigorous security definition and rigorous efficiency requirements of date peta scale storage systems. To solve the problem of using a small hash values has a proxy for the entire file, design a solution where a client proofs to the server that it indeed has the file. A proof mechanism that prevents such leakage amplification.

III. CONCLUSION

This paper compares many duplication techniques and models about the de-duplication proofs of security ,current and past history of the de-duplication to avoid the duplicate copies of the users and also provide the security. A hybrid cloud approach are used to provide the security and bandwidth .The survey paper specifies the de-duplication and the effectiveness with the security are described from the above mentioned papers. The techniques provide the users to protect the security from unauthorized users**.**

REFERENCES

[1] P. Anderson and L. Zhang. Fast and secure laptop backups with encrypted de-duplication. In *Proc. of USENIX LISA*, 2010.
[2] Bellare, C. Namprempre, and G. Neven. Security proofs for identity-based identification and signature schemes. *J. Cryptology*, 22(1):1–61, 2009.
[3] 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
[4] 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. Ownership in remote storage systems.
[5] n Y. Chen, G. Danezis, and V. Shmatikov, editors, *ACM Conference on Computer and Communications Security*, pages 491–500. ACM, 2011.

[6] J. Li, X. Chen, M. Li, J. Li, P. Lee, and W. Lou. Secure deduplication with efficient and reliable convergent In *IEEE Transactions on Parallel and Distributed Systems*, 2013.
[7] libcurl. http://curl.haxx.se/libcurl/.
[8] Ng and P. Lee. Revdedup: A reverse deduplication storage system optimized for reads to latest backups. In *Proc. of APSYS*, Apr 2013.
[9] W. K. Ng, Y. Wen, and H. Zhu. Private data deduplication protocols in cloud storage. In S. Ossowski and P. Lecca, editors, *Proceedings of the 27th Annual ACM Symposium on Applied Computing*, pages 441–446. ACM, 2012.
[10] 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.
[11] J. Stanek, A. Sorniotti, E. Androulaki, and L. Kencl. A secure data deduplication scheme for cloud storage. In *Technical Report*, 2013.
[12] M. W. Storer, K. Greenan, D. D. E. Long, and E. L. Miller. Secure data deduplication. In *Proc. of StorageSS*, 2008.
[13] Jin li,yan kit li,xiaofeng chen,patrick p.c.lee,wenjing lou.A hybrid cloud approach for secure authorized deduplication.2014.