International Journal of Engineering & Technology, 7 (2.32) (2018) 459-460

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Secure Data Deduplication by Using Convergent Key Technique

Mr.B.Tirapathi Reddy[1], Ms.K.V. Padmasree[2], Mr.K KrishnaKanth[3], Mr.DushyanthKumarReddy[4]

[1,2,3,4]Department of CSE, K L E F, Vaddeswaram, Guntur, Andhra Pradesh,India- 522502.
**Corresponding author E-mail: tirapathireddyb@kluniversity.in*

Abstract

Data de-duplication is the significant system used for removal the repetitive copy of information and broadly use in cloud storage to decrease and upload bandwidth. Our main motto be toward accomplish safe de-duplication within cloud. Convergent encryption has been utilized for safe de-duplication, complication is having colossal number of convergent keys. We will be utilizing base line approach in which each of the user have their own masterkey for encrypting their convergent keys also send them to cloud. This approach produces huge amount of keys with increment in amount of user and it needs help of user to look after master keys. At ending we will have De-key in which users need not deal with the key on their individual except steadily circulate convergent key among many servers.

Keywords:*Data de-duplication, convergent encryption, key managing, convergent key, data.*

1. Introduction

The cloud storage system to expand data storage to cloud providers. One important challenge is to make data management scalable. Storing enormous copies of similar data leads to data duplication. De-duplication removes repeating data through maintenance simply single print and referring extra repeating data to copy. Every copy have various levels of detail in a set of data it may refer to complete file or fixed-length or variable-length. In this literature, the file is duplicate it will delete the complete file or it will divide it into blocks and it will check each and every block if any block is duplicate then the block will be deleted.

Convergent encryption ensures data confidentiality. It encrypts or decrypts data print by convergent key for calculating hash tag of data print. Subsequent toward key generation with encryption, a user have the keys and sends the cipher text to the cloud.

The same data copy will produce identical convergent key with cipher text. This is utilized for de-duplication on cipher text and it can be decoded by owner through their convergent keys.

In convergent encryption we regard as base line approach. Original copy of data be primarily encrypted by means of the key that is resultant by means of print with key be next encrypted with a master key that will be there put away by the user. Encrypted keys are put away through encrypted data in cloud.The master key is use to retrieve keys and the files. In the base line approach we are having two issues. First, resolve huge amount of keys by more amount of users. Next, but master key be gone by chance user data cannot be present if identified to attacker user information will be disclosed. Because we want to construct Dekey which builds ramp secrete sharing scheme to distribute convergent keys to multiple servers.

2. Literature Review

Side channels in cloud services, the case of de-duplication in cloud storage. Deduplication have savings in both capacity and bandwidth, we should provide high privacy guarantees and reducing bandwidth. Goal must be ensure clients that the data remain private by uploading the data.

User Authentication for Cloud Computing.

In this we will provide cloud storage security mechanism.This will have architecture with three parties namely cloud service provider,third party auditor and cloud client. This security mechanism have public auditability.

This verification service can be used by all users.The third party auditor is responsible to all transactions for verification. In this we have two challenges resolved. To support multiple verification at a time and to support on demand block verification.

3. Proposed Work:

De-duplication is used to delete duplicate copies to storage space and upload bandwidth. We have two different approaches file-level de-duplication in addition to block level de-duplication.

File level deduplication: In this deduplication entire box file will be deleted.

Block level deduplication: In this deduplication it will divide into small blocks each and every block has convergent keys and tags. The tag is first sent to server it will verify same data is present or not. If present block will be deleted, if not the tag and data is stored on server.

Convergent Encryption: It encrypts/decrypts a data copy with key for computing hash value of information itself, After input generation and encryption user has keys then send to the cipher-text.This is used to perform de-duplication on ciphertext,they can only decrypted by owner with their keys.

Base Line Approach: This implements convergent encryption based on layered approach. Data print be encrypted through key and key be encrypted with master key that will be present steadily with user key be store through encrypted copy within cloud storage

3.1. Limitations

In this we are having two issues.

This will have more amount of key with increase amount of user. Master key be missing information cannot be present. If known to attacker next information will be leak.

To overcome limitations we will have original creation call Dekey. Which provide competence for both user with storage. In this primary customer who uploads data be necessary to calculate with distributes secrets share. User who own similar data need not calculate and store. To get better copy customer should have least amount of key servers. This reduces the storage of keys and dependable against failures and attacks.

Dekey: Dekey provide efficient convergent key during convergent key de-duplication with screte sharing. It supports together file level and block level deduplication. A dekey remains protected and controls a partial amount of key servers. We execute Dekey via Ramp secrete sharing. That enable key managing to adapt various confidentiality levels. It will limit transparency during upload and download operation within cloud.

3.2. Symmetric Encryption

These encryption use common secrete key k, to encrypt/decrypt in order. It consists of three primitives.

1. KeyGen(1λ)->k be key generation alogorithm to produce k by parameters.
2. Encrypt (k,M)->C be symmetric encryption to take secret k and M next output the cipher text C.
3. Decrypt (k,C)->M be symmetric decryption algorithm.

3.3. Convergent Encryption

Convergent encryption give information privacy within de-duplication. customer obtain convergent key as of print with encrypts print throughout key. User gain a tag used for duplicate, tag determination be there use to recognize duplicate. If two data copy be like after that their tag be same. For recognition user primary send tag to server to verify if the print have be previously stored. Together key with tag be derived and the tag cannot be use to include convergent key. Both the tag and encrypted information copy will be store in server.

1. KeyGen(M)->K be key generation that map a data print M to convergent key K.
2. Encrypt(K,M)->C be the symmertric encryption that take together key K and data copy M the same as inputs and next output a cipher text C.
3. TagGen(M)->T(M) be a tag generation that map data copy M and output a tag T(M).

   T(M)=TagGen(C),C=Encrypt(K,M).

3.4 Ramp Secret Sharing:

Dekey use ramp secret sharing scheme in the direction of store up convergent keys,(n,k,r)-RSSS where n>k>r>0 generate n share since secret, it be capable of improved commencing k share however cannot be there there as of less than k share along with no in order on the top secret be able to deduce from any r share.

RSSS build two function

1. Divide a unnamed S into (k-r)piece of the same amount, produce r accidental piece of the similar amount, with encodes the k piece by a non-systematic into n share of similar size
2. Improve take one k not in n share as input with next output unique secret S.

3.5 Problem Formulation:

There are three unusual entities to deal out model used by Dekey.

4. A user who wants to contract out data storage to the server and contact information later on. To keep upload bandwidth, user single uploads single information but do not upload any second copy information, which might be own through the similar customer or dissimilar user.
5. Server provide information service with supplies information on top of user. To reduce storage price, server eliminate storage space of unnecessary data and keep single single data.
6. A KM-CSP maintain convergent key for user, with provide user

With least storage and calculation services to ease key management.

3.6 File Upload

1. User upload folder F. It perform file level deduplication .On top of contribution file F, user compute with send file label to server.
2. Upon getting tag, the S-CSP check whether present exist similar tag on server. If server reply file is duplicate or not.
3. If the user receive the reply ‘‘no box file second copy next skip toward block level de-duplication. If reply be file duplicate next the user run PoWF on F by server to show that it really own similar file F to store up on server.
4. PoWF be established, server it proceed a folder indicator F in the direction of the user, and no more information will be uploaded. If PoWF fail, server abort upload process.

The user performs block level de-duplication on the way to more remove some unnecessary block.

1. On top of input case F and master key , the users perform subsequently calculation: Split F into set of block {B} for every block Bi, calculate block tag T(B)=TagGen(B) with Send set of block tag {T(B)} to server used for replacement verify.
2. Ahead getting block tags {T(B)} the server compute a block sign vector, for each i, if present exist a few store block tag to equal {T(B)}, server set 1 to show block duplicate or else it set 0 to show no block duplicate and stores {T(B)} in its storage space. After that, server profits sign vector to the user.
3. Ahead getting the signal vector, users perform next operation: for each i, if sign vector is 1, user runs PoWB on top of Bi by server to confirm to it own the block Bi. If it be accepted, server only proceeds block pointer Bi to customer Then, customer keep block pointer of Bi and do not require to upload

Bi; or else it compute encrypted block Ci =Encrypt(K,B) with convergent key K= KeyGen(B).

4. For every block B, user also compute encrypted convergent key (CK), where CK= Encrypt(k,K) with master key k also convergent key K.
5. Client uploads single block B by 0, all encrypted convergent key (CK) and T(F) to server, which next supplies them in file storage space system.

4. Conclusion

Dekey is an proficient and dependable convergent key managing system designed for secure deduplication. Dekey apply deduplication with convergent key and distributes convergent key share all through various key servers, although preserve semantic safety measures of convergent key along with privacy of outsourced information. We apply Dekey with Ramp secret sharing scheme and to gain small encoding/decoding in the cloud compare to network communication in cloud in normal upload/download operation.

References

[1] D. Harnik, B. Pinkas, and A. Shulman-Peleg, ‘‘Side Channels in Cloud Services: Deduplication in Cloud Storage,’’ IEEE Security Privacy, vol. 8, no. 6, pp. 40-47, Nov./Dec. 2010.

[2] W.K. Ng, Y. Wen, and H. Zhu, ‘‘PrivateData Deduplication Protocols in CloudStorage,’’ in Proc. 27th Annu. ACM Symp.Appl. Comput., S. Ossowski and P. Lecca, Eds., 2012, pp. 441-446.

[3] Q. Wang, C. Wang, K. Ren, W. Lou, andJ. Li, ‘‘Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing,’’ IEEE Trans. Parallel Distrib. Syst., vol. 22, no. 5, pp. 847-859, May 2011.

[4] Y. Tang, P.P. Lee, J.C. Lui, and R.Perlman, ‘‘Secure Overlay Cloud Storage with Access Control and Assured Deletion,’’IEEE Trans. Dependable Secure Comput., vol. 9, no. 6, pp. 903-916,Nov./Dec. 2012.

[5] A.D. Santis and B. Masucci, ‘‘Multiple Ramp Schemes,’’ IEEE Trans. Inf. Theory, vol. 45, no. 5, pp. 1720-1728, July 1999.

[6] J.S. Plank and L. Xu, ‘‘OptimizingCauchy Reed-Solomon Codes for Fault-Tolerant Network Storage Applications,’’ in Proc. 5th IEEE Int’l Symp. NCA,Cambridge, MA, USA, July 2006, pp. 173-180.

[7] W.K. Ng, Y. Wen, and H. Zhu, ‘‘Private Data Deduplication Protocols in Cloud Storage,’’ in Proc. 27th Annu. ACM Symp.Appl. Comput., S. Ossowski and P. Lecca, Eds., 2012, pp. 441-446.

[8] Jin Li, Xiaofeng Chen, Mingqiang Li, Jingwei Li, Patrick P.C. Lee, and Wenjing Lou IEEE Transaction on parallel and distrubuted systems, vol. 25, no. 6, june 2014.

[9] 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 Proc. 3rd Int’l Workshop Security Cloud Computing., 2011, pp. 160-167.

[10] G. Wallace, F. Douglis, H. Qian, P.Shilane, S. Smaldone, M. Cham- ness, and W. Hsu., “Characteristics of Backup Workloads in Production Systems” in Proc. 10th USENIX Conf. FAST, 2012, pp. 1-16..

[11] W. Wang, Z. Li, R. Owens, and B. Bhargava, ”Secure and Efficient Access to Outsourced Data” in Proc. CM CCSW, Nov. 2009, pp. 55-66.

.