A Strategic Review on Deduplication and Authorization in Hybrid Cloud Approach

Ganesh K[1], Arjun K[2]
M Tech[1], Assistant Professor[2]
Dept of CSE, Bheema Institute of Technology & Science, Adoni, AP, India[1&2]

ABSTRACT:
Data reduplications to eliminate duplicate copies of data, which is one of the most important and common Data compression techniques to reduce the amount of storage space to save bandwidth and cloud storage is used. To protect the privacy of sensitive data deduplication supporting convergent encryption techniques to encrypt the data before it has been suggested outside. To better protect the data security, the paper officially included in the first attempt to solve the problem of official data deduplication. Unlike traditional deduplication system, users with different privileges are considered more duplicate data outside his control. Also available are many new buildings hybrid cloud data deduplication architecture that supports dual control authority. Security analysis suggested security model proposed plan shows that it is safe in terms of the definition.

***KEYWORDS*:** Reduplication, authorized duplicate check, confidentiality, hybrid cloud, secure

1. INTRODUCTION

Cloud computing platforms and applications while hiding the details, apparently as a service to users all over the Internet unlimited "virtualized" provide resources. Today's cloud service providers offer a relatively low cost with highly available and massively parallel storage and computational resources. Cloud computing becomes more prevalent; an increasing amount of data stored in the cloud and can define access rights to the stored data being shared by users who have the specified privilege. Cloud storage service management is one of the critical challenges of ever-increasing volume of data. Scalable cloud computing to data management, data deduplication has become a well-known technique and has attracted more attention lately.

To eliminate repeating copies of the data deduplication storage is a special data compression technique. Techniques used to increase utilization of storage and data transfer networks can also be applied to reduce the number of bytes need to be sent. Instead of keeping multiple copies of the same data content, data deduplication to maintain a single physical copy and eliminates redundant data by sending this copy to the other unnecessary data. File-level or block-level deduplication or may occur. For file-level deduplication, it eliminates copies of the same file. Deduplication also eliminates duplicate files are not the same square block-level data blocks can be realized.

Despite the many benefits of data deduplication makes users vulnerable to attack sensitive data inside and outside, is emerging security and privacy concerns. Data privacy while traditional encryption is incompatible with data deduplication. In particular, traditional encryption requires different users to encrypt data with their own keys. Therefore, it will be different users different ways of encrypted copies of the same data deduplication impossible. [7] Convergent encryption, deduplication has been proposed making it possible to enforce data privacy. This / data copy data obtained by a convergent key cryptographic hash calculation of the contents, copy decipher passwords. After key generation and key data encryption to protect users and send encrypted cloud. Since the encryption process is derived from deterministic and data content, copies of the same data would produce the same key and therefore convergent same encrypted. To prevent unauthorized access, the proof of ownership of a secure protocol to duplicate user is required to provide evidence of actually having the same file. Proof of a pointer to the next user with the same file from the server without the need to upload the same file will be provided. A user can simply download their owners by convergent key related data from encrypted files can be decrypted with the pointer servers. Thus, it allows you to perform on convergent encrypted deduplication and cloud encryption prevents unauthorized users to gain access to file proof of ownership.

However, in many applications the previous authorization differential deduplication system is important to double check, I cannot see. In such a competent deduplication system, it is given a number of privileges for each user of the system during startup. All files uploaded in the cloud allowing users to perform repetitive inspections and to access files that are restricted by a series of privileges to specify what kind. Some file for its user input before sending the check duplicate user demand that a copy of this file is stored as it is possible to find a copy of this file and is paired privilege and only if you need to get this file and their privileges cloud. For example, a company will be allocated to employees in many different privileges. To save cost and efficient management, privilege and deduplication techniques to be applied to specified data to store only one copy of the same file public cloud storage server provider (SCSP) will be moved. Because privacy into account, some of the files are encrypted and access control to perform double-checked by the employees will be allowed to have the specified privilege. Convergent traditional encryption-based deduplication systems, despite the provision of a degree of privacy, does not support double-check with the differential privileges. In other words, no encryption technique based on differential concession converging deduplication is considered. At the same time you want to perform both deduplication and provisioning differential seems to contradict double check.

2. REVIEW ON DIFFERENT APPROACHES

A. CONVERGENT ENCRYPTION

Convergent encryption, deduplication of data privacy. A user (or data owner) derives a convergent key data replication and copy all original data convergence with key passwords. In addition, users also used to determine tag pairs such that obtains a label copy data. Here, label accuracy feature that keeps two copies of the data are the same, then I suppose it's the same in their labels. To identify pairs of identical copies of the first users already send server-side label to check whether stored. Convergent and labels used to understand key data privacy both independently obtained and convergent label switch and reconciliation note. Both are stored in the encrypted data replication and server side corresponding label. Keygen by (M): Officially, convergent encryption scheme, identified by their four primitive functions! Copy map data for key generation algorithm converges key KM K; Enc of (M, M)! C convergent key K and M as input data and receive a copy of the encrypted symmetric encryption algorithm, and the C output; Decca (F, C)! Mr C and is then encrypted and decryption algorithm converges key K as inputs and outputs copy of the original data M; and TagGen (M)! T (M), maps and a copy of the original data M label T (M) data algorithm that tag generation.

B. IBI & IBS SCHEMES

In this article, author, security proofs, or direct or indirect defined identity-based authentication and signature schemes provided as numerous attacks on this basis on the one hand a framework that helps to discover these programs and how to obtain and provide to the other side. In this article, the author IBI (identity-based authentication) scheme and IBS (identity-based signatures) said that an authority public key and the master secret key that contains the author. We discussed about the IB scheme. With a secret key based on this authorization may give a user identity. In case of IBS scheme, similar user identifies him and just needs signing and signature outside the control of the main public key identification information that we expect to sign the message.

C. PROOF OF OWNERSHIP FOR DE-DUPLICATION

In this article, the author describes deduplication techniques. This service is used to reduce required storage by the provider and is based on intuition. The same content will be stored for several reasons. Therefore, it is sufficient to store only a single copy. Avoiding reporting with data deduplication throat several times for this. Large data sets often exhibit high reluctance. In this paper, new security protocols are available to implement the ownership proof.

D. ARCHIVAL STORAGE

In this paper, the authors of the so-called descriptor blocks Venti .Each block network storage and content describes each stirring. The repeated copying and storage consumption is reduced. Storage application called Venti building block of construction types and this explains the design and implementation of an archival storage. The main objective is to provide Venti writing archive storage times and this is shared by many client machines and applications. Venti is a block-level network storage system designed to archive data. VAC application to store collected with a flute and directories as a single object that is used in the content of the selected data is stored as a tree blocks the vent server. Vicki model -one time writing blogs .so vent a separate collection of duplicate files each author makes a useful storage application and vents copies of a block.

E. CLOUD BACKUP, DELETION AND VERSION CONTROL

This version fading .the ensure security layer on top of today's cloud storage service for secure cloud backup system for "Fade Version" offers STD version controlled backup that is why this paper describes the authors cloud storage and secure backup system concepts Amazon S3 process is thus applied in this article to back up the various versions remove unnecessary data between design shows the cross fade version of storage. Sure does not support the deletion adds burden on the poor performance of traditional cloud backup services? Is one of the service model out of the cloud computing system and most of the businesses .the remote system software or hardware, and the case may require network computing and storage classes to obtain this data will provide resources deleting .Assured to provide cloud delivery customers have their own data backups on reliable request. Data backup fade version release of version control provides assured data backup performance with the fans that additional storage of cryptographic keys. This text is stored in the key and reducing the number of future business to be managed.

F. RevDedup

In this article the author deduplication VM image storage double, though the virtual machine eliminates (VM) said images is an important issue in backup storage growing virtualization environments for increased volume will be introduced fragmentation reading this fall sequentially performance. So to overcome this, the author proposes an end deduplication system that optimizes the VM reads RevDedup. Image backup with a new idea called reverse deduplication. This fragmentation RevDedup shifts, keeping the layout of new data on old data and old data removes duplicate.

G. ROLE BASED ACCESS CONTROL MODELS

In this article, the author RBAC (Role Based Access Control) explains. Permissions are associated with roles and role members. Roles are related to access control of users. Multi-user computer system welded. This greatly facilitates management permissions. Roles are used in different business functions assigned by the user based on their qualifications. The role is usually determined to be changed less frequently. Created to perform a particular task role. RBAC security management and collection purpose. The use of administrative roles in modern networks, OS found. Research in this area has developed a systematic approach to design problems. That future business requires a systematic methodology and an integrated approach to integrated solutions, this deficit is to develop analysis and so many restrictions.

3. CONCLUSION

In this paper, the investigation is based on official data deduplication concepts have been proposed to protect the security of data with duplicate control of users with different privileges. We also duplicate and-shoot keys with special markers of files that are created by the private cloud server hybrid cloud architecture dual control, authority supports many new deduplication construction were discussed. Security analysis inside our order specified in the proposed security model and shows that it is safe in terms of external attacks. We integrity check by assigning highly competent people deduplication plans to develop the strategy.

REFERENCES

1. P. Anderson and L. Zhang. Fast and secure laptop backups with encrypted de-duplication. In Proc. of USENIX LISA, 2010.
2. M. Bellare, S. Keelveedhi, and T. Ristenpart. Dupless: Serveraided encryption for deduplicated storage. In USENIX Security Symposium, 2013.
3. M. Bellare, S. Keelveedhi, and T. Ristenpart. Message-locked encryption and secure deduplication. In EUROCRYPT, pages 296–312, 2013.
4. M. Bellare, C. Namprempre, and G. Neven. Security proofs for identity-based identification and signature schemes. J. Cryptology, 22(1):1–61, 2009.
5. 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.
6. 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.
7. 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.
8. D. Ferraiolo and R. Kuhn. Role-based access controls. In 15th NIST-NCSC National Computer Security Conf., 1992.
9. S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg. Proofs of ownership in remote storage systems. In Y. Chen, G. Danezis, and V. Shmatikov, editors, ACM Conference on Computer and Communications Security, pages 491–500. ACM, 2011.
10. J. Li, X. Chen, M. Li, J. Li, P. Lee, andW. Lou. Secure deduplication with efficient and reliable convergent key management. In IEEETransactions on Parallel and Distributed Systems, 2013.
11. C. Ng and P. Lee. Revdedup: A reverse deduplication storage system optimized for reads to latest backups. In Proc. of APSYS, Apr 2013.
12. 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.
13. R. D. Pietro and A. Sorniotti. Boosting efficiency and security in proof of ownership for deduplication. In H. Y. Youm and Y. Won, editors, ACM Symposium on Information,

Computer and Communications Security, pages 81–82. ACM, 2012.

14. S. Quinlan and S. Dorward. Venti: a new approach to archival storage. In Proc. USENIX FAST, Jan 2002.
15. 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.
16. R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman. Role-based access control models. IEEE Computer, 29:38–47, Feb 1996.
17. J. Stanek, A. Sorniotti, E. Androulaki, and L. Kencl. A secure data deduplication scheme for cloud storage. In Technical Report, 2013.
18. M. W. Storer, K. Greenan, D. D. E. Long, and E. L. Miller. Secure data deduplication. In Proc. of StorageSS, 2008.
19. Z. Wilcox-O'Hearn and B. Warner. Tahoe: the least-authority filesystem. In Proc. of ACM StorageSS, 2008.
20. J. Xu, E.-C. Chang, and J. Zhou. Weak leakage-resilient client-side deduplication of encrypted data in cloud storage. In ASIACCS, pages 195–206, 2013.
21. J. Yuan and S. Yu. Secure and constant cost public cloud storage auditing with deduplication. IACR Cryptology ePrint Archive, 2013:149, 2013.
22. K. Zhang, X. Zhou, Y. Chen, X.Wang, and Y. Ruan. Sedic: privacyaware data intensive computing on hybrid clouds. In Proceedings of the 18th ACM conference on Computer and communications security, CCS'11, pages 515–526, New York, NY, USA, 2011. ACM.
23. Li, Jin, Y. Li, Xiaofeng Chen, Patrick PC Lee, and Wenjing Lou. "A Hybrid Cloud Approach for Secure Authorized Deduplication.", 2014

AUTHOR

Ganesh K B.Tech degree in Computer Science & Engineering in the year 2012 from JNTU Anantapur. Currently his pursuing M.Tech in Computer Science & Engineering from JNTU Anantapur. His Research and area of interest is cloud computing.

Arjun K received B.Tech degree in Computer Science & Engineering from JNTU Anantapur, India. He is awarded with M-Tech degree in Computer Science & Engineering and he guided many academic projects for the last 7 years of teaching experience. His research interests are in the field of cloud computing.