International Journal for Modern Trends in Science and Technology, 8(07): 320-325, 2022
Copyright © 2022 International Journal for Modern Trends in Science and Technology
ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0807048
Available online at: http://www.ijmtst.com/vol8issue07.html

Secure Deduplicated Cloud Storage with Encrypted Two-Party Interactions in CCPS

E. Aruna[1] | Kancharla Deepthi[2] | Vuttunoori Bhavana[2] | Badikala Roshan[2]

[1]Assistant Professor, Department of IT, Teegala Krishna Reddy Engineering College, Hyderabad, India.
[2]Department of IT, Teegala Krishna Reddy Engineering College, Hyderabad, India.

To Cite this Article
E. Aruna, Kancharla Deepthi, Vuttunoori Bhavana and Badikala Roshan. Secure Deduplicated Cloud Storage with Encrypted Two-Party Interactions in CCPS. International Journal for Modern Trends in Science and Technology 2022, 8(07), pp. 320-325. https://doi.org/10.46501/IJMTST0807048

Article Info
Received: 21 June 2022; Accepted: 22 July 2022; Published: 27 July 2022.

ABSTRACT

Cloud envisioned Cyber-Physical Systems (CCPS) is a practical technology that relies on the interaction among cyber elements like mobile users to transfer data in cloud computing. . In CCPS, cloud storage applies data deduplication techniques aiming to save data storage and bandwidth for real-time services. In this infrastructure, data deduplication eliminates duplicate data to increase the performance of the CCPS application. However, it incurs security threats and privacy risks. So, we propose a message Lock Encryption with neVer-decrypt homomorphic EncRyption (LEVER) protocol between the uploading CCPS user and cloud storage to reconcile the encryption and data deduplication.Interestingly, LEVER is the first brute-force resilient encrypted deduplication with only cryptographic two-party interactions. We perform several numerical analysis of LEVER and confirm that it provides high performance and practicality compared to theliterature

1. INTRODUCTION

The amount of data to be stored by cloud storage systems increases extremely fast. It is thus of utmost importance for Cloud Storage Providers (CSPs) to dramatically reduce the cost to store all the created data. A promising approach to achieve this objective is through data deduplication. Put simply, data deduplication keeps a single copy of repeated data. When a client wishes to store some piece of data, and if a copy of this data has already been saved in the storage system, then solely a reference to this existing copy is stored at the storage server. No duplicate is created.

Data deduplication also improves users experience by saving network bandwidth and reducing backup time when the clients perform the deduplication before uploading data to the storage server. This form of deduplication is termed as client-side deduplication, and when it is handled by the storage server it is called server-side deduplication. Due to its straightforward economical advantages, data deduplication is gaining popularity in both commercial and research storage systems.

1. 1 Problem Definition

Deduplication, security issues leading to information leakage to malicious clients. To maintain confidentiality, data integrity and data deduplication became major issues over data storage. Suffering from a lack of security, high performance, and applicability.

Data access to the third-party like cloud providers. Malware attacks and intruders were increasing to indulge in to the data server and this may lead to data loss and also confidentiality may lost.

1.2 Objective of Project

The objective of this project is to tackle the brute-force attack in cloud storage, applying in CPS. To fill the above gaps and refer to the paper's contribution, some questions arise: i) Is it possible to design an efficient encrypted data deduplication algorithm in CPS? ii) Can we assure that the proposed algorithm could bring data privacy, high/practical performance compared to the literature? And, iii) How can the encrypted method support two-party interaction between the uploader and the cloud server?

2. LITERATURE SURVEY

1) *LEVER: Secure Deduplication Cloud storage With Encrypted Two-Party Interactions In Cyber-Physical Systems.*

Title: "Understanding data deduplication ratios"

Author: M. Dutch

Description: Data deduplication and other methods of reducing storage consumption play a vital role in affordably managing today's explosive growth of data. Optimizing the use of storage is part of a broader strategy to provide an efficient information infrastructure that is responsive to dynamic business requirements. This paper will explore the significance of deduplication ratios related to specific capacity optimization techniques within the context of information lifecycle management.

2) *LEVER: Secure Deduplication Cloud storage With Encrypted Two-Party Interactions In Cyber-Physical Systems.*

Title: A study of practical deduplication

Author: D. T. Meyer and W. J. Bolosk

Description: We found that whole-file deduplication achieves about three quarters of the space savings of the most aggres- sive block-level deduplication for storage of live file systems, and 87% of the savings for backup images. We also studied file fragmentation finding that it is not prevalent, and updated prior file system metadata studies, finding that the distribution of file sizes continues to skew toward very large unstructured files.

3) *LEVER: Secure Deduplication Cloud storage With Encrypted Two-Party InteractionsIn Cyber-Physical Systems.*

Title: "Proofs of ownership in remote storage systems,

Author: S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg,

Description: In this work we identify attacks that exploit client-side deduplication, allowing an attacker to gain access to arbitrary-size files of other users based on a very small hash signatures of these files. More specifically, an attacker who knows the hash signature of a file can convince the storage service that it owns that file, hence the server lets the attacker download the entire file. (In parallel to our work, a subset of these attacks were recently introduced in the wild with respect to the Dropbox file synchronization service.)

4) *LEVER: Secure Deduplication Cloud storage With Encrypted Two-Party Interactions In Cyber-Physical Systems.*

Title: Side channels in cloud services: Deduplication in cloud storage,"

Author: D. Harnik, B. Pinkas, and A. Shulman-Peleg

Description: Cloud storage refers to scalable and elastic storage capabilities delivered as a service using Internet technologies with elastic provisioning and use based pricing that doesn't penalize users for changing their storageconsumption without notice.

5) *LEVER: Secure Deduplication Cloud storage With Encrypted Two-Party Interactions In Cyber-Physical Systems.*

Title: Dark clouds on the horizon: using cloud storage as attack vector and online slack space,"

Author: M. Mulazzani, S. Schrittwieser, M. Leithner, M. Huber, and E. Weippl

Description: Within this paper we give an overview of existing file storage services and examine Dropbox, an advanced file storage solution, in depth. We analyze the Dropbox client software as well as its transmission protocol, show weaknesses and outline possible attack vectors against users. Based on our results we show that Dropbox is used to store copyright-protected files from a popular file sharing network. Further more Dropbox can be exploited to hide files in the cloud with unlimited stor-age capacity.

3. EXISTING SYSTEM

Data deduplication keeps a single copy of repeated data. When a client wishes to store some piece of data, and if a copy of this data has already been saved in the storage system, then solely a reference to this existing copy is stored at the storage server. No duplicate is created. There are diverse forms of data deduplication. It can be done by a client solely on the data he/she has previously stored in the system, a technique commonly called intra-user deduplication, or it can be achieved by taking into account the data previously stored by all the clients. In this case it is designated as inter-user deduplication. Data deduplication also improves users experience by saving network bandwidth and reducing backup time when the clients perform the deduplication before uploading data to the storage server.

This form of deduplication is termed as client-side deduplication, and when it is handled by the storage server it is called server-side deduplication. Due to its straightforward economical advantages, data deduplication is gaining popularity in both commercial and research storage systems. several works have recently revealed important security issues leading to information leakage to malicious clients. These security concerns arise especially in systems performing an inter-user and client-side deduplication which is unfortunately the kind of deduplication that provides the best savings in terms of network bandwidth and storagespace

4. PROPOSED SYSTEM

Here, we present our message Lock ncryption with ne Verdecrypt homomorphic EncRyption (LEVER).

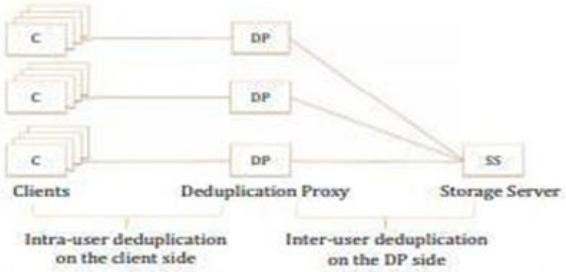
We divide the encrypted data deduplication into three phases; the user derives the chunk key for the chunk to be uploaded in the first phase, transforms the chunk to be uploaded into an encrypted form in the second phase, and then runs the ordinary data deduplication protocol in the third phase.

The design challenge of encrypted data deduplication is that a high min-entropy key can only encrypt the chunk. However, it is still difficult for different users with the same f to calculate the standard high min-entropy key.

Our work takes place in the context of an Internet Service Provider (ISP) providing also the storage system2. That is to say, the ISP which is also the CSP has strong economical reasons to (i) save storage space and network bandwidth as it masters all the network and storage infrastructure, and (ii) provide a secure storage service to its consumers.

Our deduplication scheme is simple and robust against the aforementioned attacks while remaining efficient in terms of storage space and bandwidth savings for both clients and the CSP. We consider data deduplication at a file level granularity but our solution can be extended to the block level. Specifically our approach is a two- phase deduplication that leverages and combines both intra- and inter-user deduplication techniques by introducing deduplication proxies (DPs) between the clients and the storage server (SS). Communications from clients go through these DPs to reach the SS which allows splitting the deduplication process

5. ARCHITECTURE


C
C
C
DP
DP
DP
SS
Clients
Deduplication Proxy
Storage Server
Intra-user deduplication on the client side
Inter-user deduplication on the DP side

1: Architecture of the two-phase deduplication scheme.

6. MODULES

6.1 Client (C): Any authenticated user accessing the storage system.

6.2 Storage Server (SS): A server in charge of storing and serving clients files. The storage server also maintains an index of all the files stored in the storage system and their owners.

6.3 Deduplication Proxy (DP): A server associated with a given number of clients. Clients communicate with the SS via their associated deduplication proxy. A deduplication proxy is involved in both the intra-user and the inter-user deduplication.

7. OBJECTIVES :

The output form of an information system should accomplish one or more of the following objectives.

- ❖ Convey information about past activities, current status or projections of the
- ❖ Future.
- ❖ Signal important events, opportunities, problems, or warnings.
- ❖ Trigger an action.
- ❖ Confirm an action.

8. IMPLEMENTATION

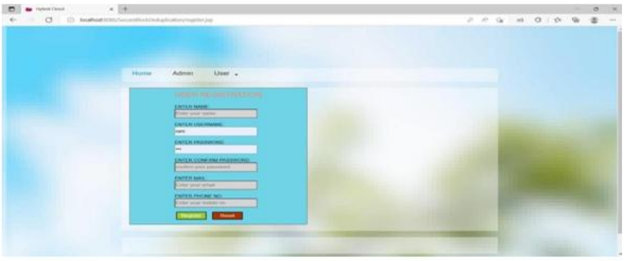
The contributions of this paper are summarized as follows. • An obstacle in designing encrypted data deduplication is how the cloud can find that two distinct encrypted chunks are from the same content. Thus, our first contribution is developing a message Lock Encryption with neVer-decrypt homomorphic EncRyption (LEVER) by taking advantage of the property of homomorphic encryption without further decryption and resiliency against brute-force by external attackers.

In LEVER, only the uploader and cloud participate in the uploading process, in contrast to most current solutions in need of the third party's participation in CPS.

LEVER can work transparently with any current cloud storage linked with the CPS users without the need for cloud storage provider's engineering work at the backend. • We validate LEVER in terms of communication and deduplication costs using two datasets, namely Enron [12] and Oxford [13].
Lastly, in addition to the analysis and numerical simulations, we have a LEVER prototype to demonstrate the practicality. A comparison among different systems is shown in Table I. The python implementation of LEVERis available in [14].

We have implemented two storage system prototypes to compare the performance overhead of our proposition with respect to a classic storage system with no data encryption. Specifically, we have developed a classic storage system with a client and storage server software modules, and one implementing our proposition with a client, a deduplication proxy and storage server software modules. All these software modules are implemented in python 2.7.6 and access the pycrypto library for the cryptographic operations. We use the SHA256 algorithm as the hash function, RSA1024 for asymmetric encryption operations, and AES for the symmetric encryption operations with keys that are 256 bits long. The storage servers use the MongoDB3 database to store the meta-data of the stored files as well as files owners. The software modules are executed on three different virtual machines (VMs) running on Ubuntu 12.04.4 LTS with 1GB of memory and an AMD Opteron(TM) octa-core Processor 6220@3GHz. The network topology is as follows: the VM executing the DP software is located on the network path between the VM running the different clients modules and the one executing the different SSs modules.

9. OUTPUT SCREENS



USER REGISTRATION

ENTER NAME: Enter your name

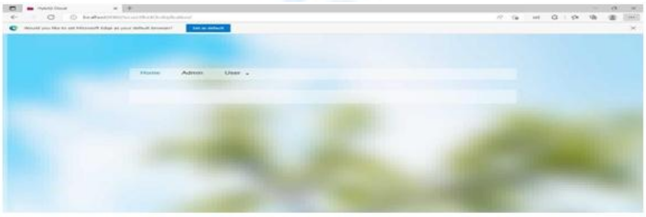
ENTER USERNAME: sam

ENTER PASSWORD: ***

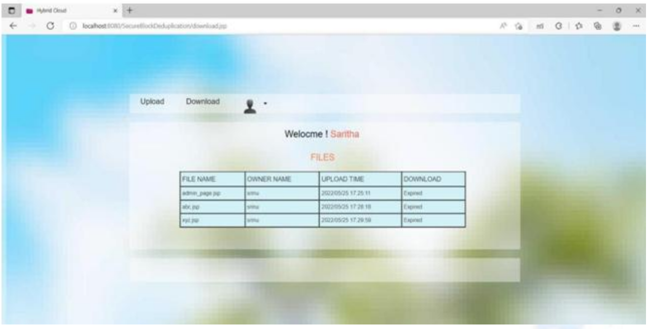
ENTER CONFIRM PASSWORD: confirm your password

ENTER MAIL: Enter your email

ENTER PHONE NO: Enter your mobile no



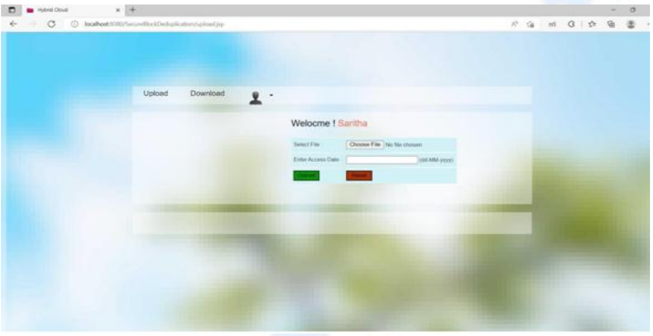
Would you like to set Microsoft Edge as your default browser?



Welocme ! Saritha

FILES

FILE NAME	OWNER NAME	UPLOAD TIME	DOWNLOAD
admin_page.jsp	srinu	2022/05/25 17:25:11	Expired
abc.jsp	srinu	2022/05/25 17:28:18	Expired
xyz.jsp	srinu	2022/05/25 17:29:59	Expired



Welocme ! Saritha

Select File: No file chosen

Enter Access Date: (dd-MM-yyyy)

10. CONCLUSION

This paper proposed a message Lock Encryption with neVer-decrypt homomorphic EncRyption (LEVER) protocol in a cloud envisioned cyber-physical system. LEVER applied encrypted client-side data deduplication with only symmetrically cryptographic two-party interactions to cope with a brute-force attack in cloud storage. We proved the security of LEVER via ideal/real paradigms. We also demonstrate the significant performance of the LEVER via the analysis and numerical simulations. Although client-side data deduplication has been widely adopted by commercial cloud storage services to eliminate data redundancy, it breaches the user privacy. Exchanging messages between the user and cloud storage to verify whether a file is already stored creates a side-channel for the attacker to gain information about the file existence status. In the future, we plan to enhance the LEVER's privacy to deal with side-channel attacks in the cloud storage by creating a probabilistic relation between the messages exchanged.

In this paper, we have presented a two-phase deduplication scheme that (i) ensures that a client actually owns the file he/she wants to store by applying an intra-user deduplication on the client side (ii) ensures that a file corresponds to its claimed identifier through a control by a deduplication proxy located between clients and the storage server and (iii) applies an inter-user deduplication on the deduplication proxy side that makes this inter-user deduplication unnoticeable to clients by adding some delay to put operations so that the length of a file upload is indistinguishable from an upload of its reference. Our method provides protection against attacks from malicious clients, global storage space savings to the CSPs thanks to the inter-user deduplication, per-client bandwidth network savings between clients and the deduplication proxies, and global network bandwidth savings between the deduplication proxies and the storage server.

For future works, we plan to address the confidentiality issues against attacks that can be performed by the CSP. We also plan to extend our solution so that encrypted decryption keys can also be deduplicated without jeopardizing security properties. We will also consider how to extend the deduplication in our schemeto a block level granularity.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

References

[1] M. Dutch, "Understanding data deduplication ratios," SNIA Data Management Forum, 2008.

[2] D. Russel, "Data deduplication will be even bigger in 2010," Gartner, February 2010.

[3] D. T. Meyer and W. J. Bolosky, "A study of practical deduplication," in Proceedings of the 9th USENIX Conference on File and Storage Technologies (FAST), 2011.

[4] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Proofs of ownership in remote storage systems," in Proceedings of the 18th ACM Conference on Computer and Communications security (CCS), 2011.

[5] D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Side channels in cloud services: Deduplication in cloud storage," IEEE Security Privacy, vol. 8, 2010.

[6] M. Mulazzani, S. Schrittwieser, M. Leithner, M. Huber, and E. Weippl, "Dark clouds on the horizon: using cloud storage as attack vector and online slack space," in Proceedings of the 20th USENIX Conference on Security (SEC), 2011.

[7] M. W. Storer, K. Greenan, D. D. Long, and E. L. Miller, "Secure data deduplication," in Proceedings of the 4th ACM International Workshop on Storage Security and Survivability (StorageSS), 2008.

[8] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Dupless: Server-aided encryption for deduplicated storage," in Proceedings of the 22nd USENIX Conference on Security (SEC), 2013.

[9] K. Suzaki, K. Iijima, T. Yagi, and C. Artho, "Memory deduplication as a threat to the guest os," in Proceedings of the 4th ACM European Workshop on System Security (EUROSEC), 2011.

[10] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Message-locked encryption and secure deduplication," in Proceedings of the 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT), 2013.

[11] R. Di Pietro and A. Sorniotti, "Boosting efficiency and security in proof of ownership for deduplication," in Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security (ASIACCS), 2012.

[12] O. Heen, C. Neumann, L. Montalvo, and S. Defrance, "Improving the resistance to side-channel attacks on cloud storage services," in Proceedings of the 5th International Conference on New Technologies, Mobility and Security (NTMS), 2012.

[13] C. Liu, X. Liu, and L. Wan, "Policy-based de-duplication in secure cloud storage," in Trustworthy Computing and Services, ser. Communications in Computer and Information Science. Springer Berlin Heidelberg, 2013, vol. 320, pp. 250–262.

[14] W. K. Ng, Y. Wen, and H. Zhu, "Private data deduplication protocols in cloud storage," in Proceedings of the 27th Annual ACM Symposium on Applied Computing (SAC), 2012.

[15] J. Xu, E.-C. Chang, and J. Zhou, "Weak leakage-resilient client-side deduplication of encrypted data in cloud storage," in Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security (ASIACCS), 2013.

[16] Q. Zheng and S. Xu, "Secure and efficient proof of storage with deduplication," in Proceedings of the 2nd ACM Conference on Data and Application Security and Privacy (CODASPY), 2012.

[17] S. Kamara and K. Lauter, "Cryptographic cloud storage," in Proceedings of the 14th International Conference on Financial Cryptography and Data Security (FC), 2010.

[18] K. Ren, C. Wang, and Q. Wang, "Security challenges for the public cloud," IEEE Internet Computing, vol. 16, 2012.

[19] F. Rocha, S. Abreu, and M. Correia, “The final frontier: Confidentiality and privacy in the cloud,” Computer, vol. 44, no. 9, pp. 44–50, 2011.

[20] 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 Proceedings of the 22nd IEEE International Conference on Distributed Computing Systems (ICDCS), 2002.

[21] G. Pallis and A. Vakali, “Insight and perspectives for content delivery networks,” Communications of the ACM, vol. 49, no. 1, pp. 101–106, 2006.

[22] C. Huang, A. Wang, J. Li, and K. W. Ross, “Understanding hybrid CDN-P2P: why Limelight needs its own Red Swoosh,” in Proceedings of the 18th ACM International Workshop on Network and Operating System support for digital audio and video (NOSSDAV), 2008.